

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

UTILITY PATENT APPLICATION FOR

**METHODS OF DIRECTING C-O BOND FORMATION UTILIZING A TYPE II
POLYKETIDE SYNTHASE SYSTEM**

Inventor(s): **BEN SHEN**, a citizen of the Peoples Republic of China, residing
at 1806 Autumn Hill Drive, Verona, Wisconsin 53593

HYUNG-JIN KWON, a citizen of South Korea, residing
at 4801 Sheboygan Avenue, Room 307, Madison, Wisconsin
53705

Assignee: Wisconsin Alumni Research Foundation

Entity: Small Entity

Claims : 25

Attorney Docket No.: 054030-0031

Drawings: 16 sheets

Sequence Listing: 18 sheets

<p>EXPRESS MAIL Mailing Label Number <u>EF126344828US</u></p> <p>Date of Deposit <u>August 22, 2003</u></p> <p>I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" Service under 37 CFR 1.10 on the date indicated above and is addressed to Mail Stop Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450</p> <p><u>Luanne Connor</u> (Typed or printed name of person mailing paper or fee)</p> <p><u>Luanne Connor</u> (Signature of person mailing paper or fee)</p>	<p>Charles L. Leeck Registration No. 50,343 Sonali S. Srivastava Registration No. 52,248</p> <p>GODFREY & KAHN, S.C. 780 North Water Street Milwaukee, Wisconsin 53202 (414) 273-3500</p>
--	---

METHODS OF DIRECTING C-O BOND FORMATION UTILIZING A
POLYKETIDE SYNTHASE

RELATED APPLICATIONS

5 This application claims the benefit of U.S. Provisional Application No.
60/405,245 filed on August 22, 2002, which is incorporated herein by reference in its
entirety for all purposes.

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY
SPONSORED RESEARCH AND DEVELOPMENT

10 This work was supported in part by a grant from the National Science Foundation
MCB-0196528. The Government of the United States of America may have certain
rights in this invention.

FIELD OF THE INVENTION

15 This invention relates to the field of polyketide biosynthesis. In particular, this
invention is directed to methods of using polyketide synthases and components thereof
capable of catalyzing C-O bonds.

BACKGROUND OF THE INVENTION

20 Polyketides are a large family of structurally diverse natural products found in
bacteria, fungi, and plants. Many polyketides are clinically important drugs such as
erythromycin and tetracycline (antibacterial), daunorubicin and epothilone (anticancer),
FK-506 and rapamycin (immunosuppressant), and lovastatin (antihypercholesterolemic).
Despite their apparent structural diversity, polyketides share a common mechanism of
biosynthesis. The carbon backbone of a polyketide results from sequential condensation

of acyl coenzyme (CoA) precursors, and this process, in a mechanistic analogy to fatty acid biosynthesis by the fatty acid synthases (FASs), is catalyzed by the polyketide synthases (PKSs). Much of the current research on polyketide biosynthesis is driven by the following two factors: (1) the extraordinary structure, mechanism, and catalytic reactivity of PKSs that provide an unprecedented opportunity to investigate the molecular mechanism of enzyme catalysis, molecular recognition, and protein-protein interaction and (2) the remarkable flexibility and plasticity of PKSs that allow the production of novel compounds that are difficult to access by traditional chemical synthesis.

The success of the biosynthetic approach depends critically on the availability of novel genetic systems and on genes encoding novel enzyme activities. Of the various polyketide pathways currently being examined, the macrotetrolides, of polyketide origin, offer a distinct opportunity to study the biosynthesis of unique molecular scaffolds. Understanding the mechanisms underlying macrotetrolide production will provide access to rational engineering of macrotetrolide analogues for novel drug leads and also facilitate construction of improved macrotetrolide expression strains by altering specific components of the biosynthetic machinery. In addition, and more importantly, elucidation of the molecular mechanisms of macrotetrolide biosynthesis contribute to the general field of combinatorial biosynthesis by expanding the repertoire of novel PKSs available for use in combinatorial biosynthesis schemes. Consequently, identifying novel PKS mechanisms will further expand the size and diversity of chemical libraries accessible by combinatorial biosynthesis.

SUMMARY OF THE INVENTION

In general, the present invention provides methods of modifying biological molecules based on the C-O bond forming activities of polypeptides derived from or related to a type II polyketide synthase (PKS) system capable of C-O bond formation. In particular, the inventors have recently identified the unique C-O bond forming activity of NonJ and NonK, two ketoacyl synthases present within the type II PKS responsible for biosynthesis of the macrotetralide nonactin. NonJ and NonK are unique in their ability to act directly on acyl-CoA intermediates in catalyzing C-O bond formation. Based on the related disclosure provided herein, one of skill in the art will realize many practical and useful advantages provided by the present invention. The amino acid sequences for NonJ and NonK are set forth in SEQ ID NOS. 3 and 5, respectively. As well, nucleic acid sequences encoding NonJ and NonK are provided in SEQ ID NOS. 2 and 4, respectively. SEQ ID NO. 1 sets forth a partial nucleic acid sequence of the nonactin biosynthesis gene cluster including both *nonJ* and *nonK* genes.

In certain embodiments, the invention provides methods including steps of contacting a biological molecule which is a substrate for a polypeptide selected from the group consisting of: (a) a polypeptide comprised by an amino acid sequence set forth in SEQ ID NO. 3; (b) a polypeptide encoded by a nucleic acid comprising nucleotide sequence set forth in SEQ ID NO. 2; and (c) a polypeptide encoded by a nucleic acid that specifically hybridizes under stringent conditions to SEQ ID NO. 2 and capable of C-O bond formation; with the polypeptide. The polypeptide then modifies the biological molecule by formation of a C-O bond.

In preferred embodiments, a method according to the invention includes the further step of contacting the biological molecule modified by the polypeptide described in the preceding paragraph with a second polypeptide selected from the group consisting of: (a) a polypeptide comprised by an amino acid sequence set forth in SEQ ID NO. 5; (b) a polypeptide encoded by a nucleic acid comprising nucleotide sequence set forth in SEQ ID NO. 4; and (c) a polypeptide encoded by a nucleic acid that specifically hybridizes under stringent conditions to SEQ ID NO. 4 and capable of C-O bond formation. The second polypeptide further modifies the biological molecule by formation of a C-O bond.

In particularly preferred embodiments, methods according to the invention facilitate C-O bond formation between the biological molecule and a second biological molecule, the second biological molecule also suitable as a substrate for the polypeptide.

In certain embodiments, the methods described and claimed herein are carried out *in vivo*, through the use of a host cell. Suitable host cells include bacterium as well as eukaryotic cells such as mammalian cells, yeast cells, plant cells, fungal cells, and insect cells. In alternate embodiments, the present invention may be practiced *ex vivo*. In yet other embodiments, at least one of the biological molecules utilized as substrate is exogenously supplied. As well, methods according to the invention preferably result in the production of a macrotetralide or a macrotetralide analogue.

In various embodiments, the invention encompasses yet additional methods of catalyzing a C-O bond between biological molecules. Such methods include steps of contacting biological molecules which are substrates for at least one polypeptide capable of catalyzing C-O bond formation between the biological molecules and encoded by a

nucleic acid set forth in SEQ ID NO. 1 or a nucleic acid hybridizing under stringent conditions thereto, with the polypeptide. The polypeptide then catalyzes C-O bond formation between the biological molecules.

In preferred embodiments, the invention provides methods of producing a
5 macrotetralide or a macrotetralide analogue. These methods include steps of contacting
biological molecules that are substrates for at least one polypeptide selected from the
group consisting of: (a) a polypeptide encoded by an amino acid sequence set forth in
SEQ ID NO. 3 or 5; (b) a polypeptide encoded by a nucleic acid comprising a nucleotide
sequence set forth in SEQ ID NO. 2 or 4; and (c) a polypeptide encoded by a nucleic acid
10 that specifically hybridizes under stringent conditions to SEQ ID NO. 2 or 4 and capable
of C-O bond formation; with the polypeptide under conditions such that the polypeptide
catalyzes a C-O bond between the biological molecules. A macrotetralide or
macrotetralide analogue is thereby synthesized and the macrotetralide or macrotetralide
analogue is subsequently recovered.

15 In other embodiments, the invention is directed to methods of preparing a hybrid
enzyme. Such methods include the step of positioning in a hybrid enzyme at least one
catalytic domain capable of catalyzing C-O bond formation between biological
molecules. The catalytic domain is encoded by a polypeptide selected from the group
consisting of: (a) a polypeptide encoded by an amino acid sequence set forth in SEQ ID
20 NO. 3 or 5; (b) a polypeptide encoded by a nucleic acid comprising nucleotide sequence
set forth in SEQ ID NO. 2 or 4; (c) a polypeptide encoded by a nucleic acid that
specifically hybridizes under stringent conditions to SEQ ID NO. 2 or 4 and capable of
C-O bond formation.

In yet other embodiments, the invention provides methods of preparing a megasynthetase. These methods include the step of positioning in a megasynthetase at least one module including a polypeptide capable of catalyzing C-O bond formation between biological molecules. The polypeptide is selected from the group consisting of:

5 (a) a polypeptide encoded by an amino acid sequence set forth in SEQ ID NO. 3 or 5; (b) a polypeptide encoded by a nucleic acid comprising nucleotide sequence set forth in SEQ ID NO. 2 or 4; and (c) a polypeptide encoded by a nucleic acid that specifically hybridizes under stringent conditions to SEQ ID NO. 2 or 4 and capable of C-O bond formation.

10 The present invention also includes methods of catalyzing C-O bond formation between biological molecules by NonJ and polypeptides derived there from. Methods according to the invention include steps of contacting biological molecules that are substrates for a polypeptide selected from the group consisting of: (a) a polypeptide comprised by an amino acid sequence set forth in SEQ ID NO. 3; (b) a polypeptide

15 encoded by a nucleic acid comprising nucleotide sequence set forth in SEQ ID NO. 2; and (c) a polypeptide encoded by a nucleic acid that specifically hybridizes under stringent conditions to SEQ ID NO. 2 and capable of C-O bond formation; with the polypeptide whereby the polypeptide catalyzes C-O bond formation between the biological molecules.

20 In particular embodiments according to the invention, methods are provided for catalyzing C-O bond formation between biological molecules by NonK and polypeptides derived there from. Such methods include steps of contacting biological molecules that are substrates for a polypeptide selected from the group consisting of: (a) a polypeptide

comprised by an amino acid sequence set forth in SEQ ID NO. 5; (b) a polypeptide encoded by a nucleic acid comprising nucleotide sequence set forth in SEQ ID NO. 4; and (c) a polypeptide encoded by a nucleic acid that specifically hybridizes under stringent conditions to SEQ ID NO. 4 and capable of C-O bond formation; with the

5 polypeptide. The polypeptide then catalyzes C-O bond formation between the biological molecules.

In yet another embodiment, the present invention provides methods of chemically modifying a biological molecule by C-O bond formation. Generally, the method includes contacting a biological molecule that is a substrate for a polypeptide selected from the

10 group consisting of: a polypeptide encoded by an amino acid sequence set forth in SEQ ID NO. 3 or 5; a polypeptide encoded by a nucleic acid comprising nucleotide sequence identical to or isolated from SEQ ID NO. 1, 2 or 4; a polypeptide encoded by a nucleic acid encoding an amino acid sequence set forth in SEQ ID NO. 3 or 5; and a polypeptide encoded by a nucleic acid that specifically hybridizes under stringent conditions to SEQ

15 ID NO. 1, 2 or 4. The particular polypeptide chemically modifies the biological molecule by formation of a C-O bond. In another embodiment of the invention, the foregoing method effectuates a condensation between the biological molecule and a second biological molecule.

These and other applications and advantages of the present invention will become

20 apparent from the detailed description accompanying the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1. Early proposed pathway for macrotetrolide biosynthesis in *S. griseus*.

Fig. 2. The macrotetrolide biosynthetic gene cluster from *S. griseus* DSM40695 (A) and HPLC analysis of macrotetrolide production in *S. lividans* (pBS2013) (B) and *S. griseus* DSM40695 (C). The region contained in pBS2013 is boxed.

Fig. 3. A restriction map of a 55 Kb contiguous DNA region from *S. griseus* DSM40695 showing general organization of the macrotetrolide biosynthetic gene cluster.

Fig. 4. Structures and mechanisms of bacterial polyketide synthases. (A) Type I PKS consists of noniteratively acting domains, (B) type II PKS consists of iteratively acting subunits, and both types utilize ACP to activate substrates and channel intermediates. (C) Type III PKS consists of iteratively acting single subunit and utilizes acyl CoA directly as substrates. Domains or subunits shown are hypothetical. “x”, “y”, and “n” emphasize the stoichiometry of subunits. The KS α and KS β subunits in solution exist as a $\alpha_2\beta_2$ heterotetramer for the actinorhodin PKS and as a $\alpha\beta$ heterodimer for the tetracenomycin PKS. The stoichiometric relationship between KS α /KS β and ACP subunits is unknown. The stoichiometry for type III PKS of bacterial origin is unknown although two type III PKS from plants have been determined to be homodimers. “z” and “m” indicates the number of cycles the starter unit is elongated by the extender units to afford the full-length polyketide products. ACP, acyl carrier protein; AT, acyl transferase; DH, dehydratase; ER, enoyl reductase; KR, ketoreductase; KS, ketoacyl synthase.

Fig. 5. C—C bond-forming step catalyzed by (A) type I and –II PKSs and (B) type III PKSs and (C) C—O bond-forming steps generally catalyzed by the NonJK KSs. KS, β -ketoacyl synthase; ACP, acyl carrier protein; CoA, coenzyme A; MCoA, malonyl CoA.

Fig. 6. (A) The macrotetrolide biosynthesis gene cluster with the five *nonJKPQU* KS genes shown in medium shading and the *nonL* CoA ligase gene shown in black. (B) Biotransformation of (\pm)-3 into 1 by *S. lividans* strains harboring various *non* gene expression cassettes. (Shown in black are genes cassettes constructed in pSET152 and shown in blue are gene cassettes constructed in pWHM3). Black rectangles with arrow indicate orientation of the *actI* promoters. Each entry represents an *S. lividans* strain transformed with either the pSET152 or pWHM3-based cassette or both and its biotransformation yield of 1 from 5 mg of (\pm)-3. The *nonK* C161G and *nonJ* C169G mutants are shown in. (C) Examples of HPLC chromatograms of biotransformation of (\pm)-3 into 1 by various *S. lividans* recombinant strains: I, authentic 1 standard (\blacklozenge); II, *S. lividans* (pBS2014/pBS2015); V, *S. lividans* (pBS2017); VIII, *S. lividans* (pBS2019/pBS2018); IX, *S. lividans* (pBS2020/pBS2018).

Fig. 7. *In vitro* synthesis of (\pm)-2 from (\pm)-3 catalyzed by NonL, requiring ATP and CoA. (A) Negative control in the absence of NonL and (B) complete assay with 800 nM NonL. ATP, (\bullet); AMP, (\blacklozenge); CoA, (\emptyset); (\pm)-2, (∇).

Fig. 8. Alignments of the conserved catalytic residues of NonJK with KSs from FAS, type I and -II PKS and type III PKS. The conserved Cys residue for C—C or C—O bond formation and His-His or His-Asn residues for decarboxylation are highlighted. Protein accession numbers are given after the protein names: FabB, P14926; FabF, P39435; DEBS1, Q03131; PikAIV, AAC69332; Act KS α , CAC44200; Tcm KS α , AAA67515; CHS2, P30074; RppA, BAA33495; NonJ, AAD37451; and NonK, AAD37450.

Fig. 9. A map of *non* gene expression cassettes as exemplified by pBS2019 for *nonJK* under the *actI* promoter. The *actI* promoter as a *HindIII*-*NdeI* fragment and the *nonJK* genes as an *EcoRI*-*NsiI* fragment were cloned into pSET152. Restriction sites: E, *EcoRI*; Ev, *EcoRV*; H, *HindIII*; N, *NdeI*; Ns, *NsiI*; P, *PacI*; Ps, *PstI*; Sp, *SpeI*; X, *XbaI*.

5 Fig. 10. SDS-PAGE analysis of *nonL* expression and NonL purification from BL21 (DE3) (pBS2023). NonL was purified by affinity chromatography on Ni-NTA resin according to the manufacturer's procedure (Novagen). Lanes: 1, molecular weight marker; 2, total lysate before IPTG induction; 3, total lysate after IPTG induction; 4-6, NonL eluted from Ni-NTA resin with 250 mM imidazole.

10 Fig. 11. Biosynthetic pathway of macrotetrolides in *S. griseus*. Arrows with broken lines indicate intermediates that have been isolated or confirmed by feeding experiments. The macrotetramerization process is catalyzed by NonL, NonJ, and NonK.

Fig. 12. (A) *In vitro* synthesis of 6-CoA and its derivatives from 6 by NonL (I), *S. lividans* (pBS2045)-CFE, negative control in the absence of 6 (II), *S. lividans* (pBS2045)-CFE (III), and *E. coli* BL21 (pBS2046)-CFE (IV). The purified NonL or CFEs were incubated with 2 mM ATP, 0.5 mM CoA, and 1 mM 6 for 1 h at 30°C. (B) *In vitro* synthesis of 6-CoA and 8-CoA by *S. lividans* (pBS2045)-CFE obtained from the nickel affinity chromatography (see the text). The CFE was incubated with 2 mM ATP, 0.5 mM CoA, and 1 mM 6 for 1h (I) and 2hs (II) at 30°C.

20 Fig. 13. Production of 1 from 6 assayed by TLC (A) and Evaporative Light-Scattering Detector (ELSD)-HPLC analysis (B) with authentic 1 and 6-CoA and 8-CoA (C) by NonL (C-I), *S. lividans* (pBS2048)-CFE (II), and *S. lividans* (pBS2049)-CFE (III),

respectively. For A, II and III were conducted as duplicates. The CFEs were incubated with 2 mM each of ATP, CoA, and 6 for 1 h at 30°C.

Fig. 14. (A) *In vitro* synthesis of 6-CoA (☒) and 8-CoA (☒) by *S. lividans* (pBS2654). NonL reaction prepared by 1 mM each of ATP, CoA, and 6 for 1 h at 30°C (I) and supplemented with the CFEs of *S. lividans* (pWHM3) (II), *S. lividans* (pBS2053) (III), and *S. lividans* (pBS2054) (IV) and then maintained for additional 1h at 30°C. (B) Formation of 6-CoA or 6-dp-CoA by *S. lividans* (pBS2054)-CFE incubated with 100 μM each of (-)-6-CoA and (-)-6-dp-CoA (I), (-)-6-CoA and (+)-6-dp-CoA (II), (+)-6-CoA and (-)-6-dp-CoA (III), for 1 h at 30°C.

Fig. 15. Evaporative Light-Scattering Detector (ELSD)-High Performance Liquid Chromatography (HPLC) analysis of authentic nonactin (I), macrotetrolides isolated from *S. griseus* DSM40695 without (II) or with the supplementation of 8 (III), *S. griseus* SB2003 with the supplementation of 8 (IV).

Fig. 16. *In vitro* assay of NonJ with (±)-6-CoA (II), (±)-8-CoA (III), (±)-6-CoA and (±)-8-CoA (IV), (+)-6-CoA and (±)-8-CoA (V), (-)-6-CoA and (±)-8-CoA (VI) and the control without NonJ (I). High Performance Liquid Chromatography (HPLC) analysis was performed on an Alltima C18 column (4.6 x 250 mm, 5 μ, 100 Å, Alltech, Deerfield, IL). The column was first eluted in 5 % CH₃CN in 20 mM NH₄Ac, pH 5.8 (buffer A) for 7 min, followed by linear gradients from 100 % buffer A to 50:50 of (buffer A:CH₃CN) for 8 min, from 50:50 of (buffer A:CH₃CN) to 10:90 of (buffer A:CH₃CN) and then maintained for additional 10 min at the flow rate of 1.0 ml/min with UV-detection at 260 nm.

DETAILED DESCRIPTION OF THE INVENTION

I. IN GENERAL

Before the present materials methods are described, it is understood that this invention is not limited to the particular methodology, protocols, cell lines, and reagents described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a cell" includes a plurality of such cells and equivalents thereof known to those skilled in the art, and so forth. As well, the terms "a" (or "an"), "one or more" and "at least one" can be used interchangeably herein. It is also to be noted that the terms "comprising", "including", "characterized by" and "having" can be used interchangeably.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are now described. All publications and patents mentioned herein are incorporated by reference for all purposes including describing and disclosing the chemicals, cell lines, vectors, animals, instruments, statistical analysis and methodologies which are reported in the publications which might be used in connection with the invention. All references cited in this specification are to be taken as indicative

of the level of skill in the art. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, for example, Molecular Cloning A Laboratory Manual, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); DNA Cloning, Volumes I and II (D. N. Glover ed., 1985); Oligonucleotide Synthesis (M. J. Gait ed., 1984); Mullis et al. U.S. Pat. No: 4,683,195; Nucleic Acid Hybridization (B. D. Hames & S. J. Higgins eds. 1984); Transcription And Translation (B. D. Hames & S. J. Higgins eds. 1984); Culture Of Animal Cells (R. I. Freshney, Alan R. Liss, Inc., 1987); Immobilized Cells And Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide To Molecular Cloning (1984); the treatise, Methods In Enzymology (Academic Press, Inc., N.Y.); Gene Transfer Vectors For Mammalian Cells (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); Methods In Enzymology, Vols. 154 and 155 (Wu et al. eds.), Immunochemical Methods In Cell And Molecular Biology (Mayer and Walker, eds., Academic Press, London, 1987); Handbook Of Experimental Immunology, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986).

II. DEFINITIONS

The terms "*nonJ*, *nonK*, *nonL*, *nonJK*, or *nonJKL* open reading frame", and "*nonJ*, *nonK*, *nonL*, *nonJK* or *nonJKL* ORF" refer to an open reading frame in the nonactin biosynthesis gene cluster as isolated from *Streptomyces griseus*. The term also embraces the same open reading frames as present in other macrotetrolide-synthesizing organisms

(e.g. other strains and/or species of *Streptomyces*, and the like). The term encompasses allelic variants and single nucleotide polymorphisms (SNPs). In certain instances the defined terms are used synonymously with the encoded polypeptide may include conservative amino acid substitutions. The term "open reading frame" shall also include
5 subsets of nucleotides as exemplified by those encoding domains encompassing a particular catalytic site of an enzyme. The particular usage will be clear from context.

The terms "isolated" "purified" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany it as found in its native state. With respect to nucleic acids and/or polypeptides the term can refer to
10 nucleic acids or polypeptides that are no longer flanked by the sequences typically flanking them in nature. A "purified" polypeptide shall also refer to a polypeptide recombinantly expressed in a host cell and utilized according to the present invention in the *in vivo* setting.

The terms "polypeptide", "peptide" and "protein" are used interchangeably herein
15 to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers. The term also includes variants on the traditional peptide linkage joining the amino acids making up the polypeptide.

20 The terms "nucleic acid" or "oligonucleotide" or grammatical equivalents herein refer to at least two nucleotides covalently linked together. A nucleic acid of the present invention is preferably single-stranded or double stranded and will generally contain phosphodiester bonds, although in some cases, as outlined below, nucleic acid analogs

are included that may have alternate backbones, comprising, for example, phosphoramidate (Beaucage *et al.* (1993) *Tetrahedron* 49:1925) and references therein; Letsinger (1970) *J. Org. Chem.* 35:3800; Sprinzl *et al.* (1977) *Eur. J. Biochem.* 81: 579; Letsinger *et al.* (1986) *Nucl. Acids Res.* 14: 3487; Sawai *et al.* (1984) *Chem. Lett.* 805, Letsinger *et al.* (1988) *J. Am. Chem. Soc.* 110: 4470; and Pauwels *et al.* (1986) *Chemica Scripta* 26: 141 9), phosphorothioate (Mag *et al.* (1991) *Nucleic Acids Res.* 19:1437; and U.S. Patent No. 5,644,048), phosphorodithioate (Briu *et al.* (1989) *J. Am. Chem. Soc.* 111 :2321, O-methylphosphoramidite linkages (see Eckstein, *Oligonucleotides and Analogues: A Practical Approach*, Oxford University Press), and peptide nucleic acid backbones and linkages (see Egholm (1992) *J. Am. Chem. Soc.* 114:1895; Meier *et al.* (1992) *Chem. Int. Ed. Engl.* 31: 1008; Nielsen (1993) *Nature*, 365: 566; Carlsson *et al.* (1996) *Nature* 380: 207). Other analog nucleic acids include those with positive backbones (Denpcy *et al.* (1995) *Proc. Natl. Acad. Sci. USA* 92: 6097; non-ionic backbones (U.S. Patent Nos. 5,386,023, 5,637,684, 5,602,240, 5,216,141 and 4,469,863; Angew. (1991) *Chem. Intl. Ed. English* 30: 423; Letsinger *et al.* (1988) *J. Am. Chem. Soc.* 110:4470; Letsinger *et al.* (1994) *Nucleoside & Nucleotide* 13:1597; Chapters 2 and 3, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook; Mesmaeker *et al.* (1994), *Bioorganic & Medicinal Chem. Lett.* 4: 395; Jeffs *et al.* (1994) *J. Biomolecular NMR* 34:17; *Tetrahedron Lett.* 37:743 (1996) and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, *Carbohydrate Modifications in Antisense Research*, Ed. Y.S. Sanghui and P. Dan Cook. Nucleic acids containing one or more carbocyclic sugars are also included within the definition of nucleic acids (see Jenkins *et*

al. (1995), *Chem. Soc. Rev.* pp169-176). Several nucleic acid analogs are described in Rawls, C & E News June 2, 1997 page 35. These modifications of the ribose-phosphate backbone may be done to facilitate the addition of additional moieties such as labels, or to increase the stability and half-life of such molecules in physiological environments.

5 The term "heterologous" as it relates to nucleic acid sequences such as coding sequences and control sequences, denotes sequences that are not normally associated with a region of a recombinant construct, and/or are not normally associated with a particular cell. Thus, a "heterologous" region of a nucleic acid construct is an identifiable segment of nucleic acid within or attached to another nucleic acid molecule that is not found in
10 association with the other molecule in nature. For example, a heterologous region of a construct could include a coding sequence flanked by sequences not found in association with the coding sequence in nature. Another example of a heterologous coding sequence is a construct where the coding sequence itself is not found in nature (e.g., synthetic sequences having codons different from the native gene). Similarly, a host cell
15 transformed with a construct which is not normally present in the host cell would be considered heterologous for purposes of this invention.

 A "coding sequence" or a sequence which "encodes" a particular polypeptide (e.g. a PKS, *etc.*), is a nucleic acid sequence which is ultimately transcribed and/or translated into that polypeptide *in vitro* and/or *in vivo* when placed under the control of appropriate
20 regulatory sequences. In certain embodiments, the boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A coding sequence can include, but is not limited to, cDNA from procaryotic or eucaryotic mRNA, genomic DNA sequences from procaryotic or

eucaryotic DNA, and even synthetic DNA sequences. In preferred embodiments, a transcription termination sequence will usually be located 3' to the coding sequence.

Expression "control sequences" refers collectively to promoter sequences, ribosome binding sites, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers, and the like, which collectively provide for the transcription and translation of a coding sequence in a host cell. Not all of these control sequences need always be present in a recombinant vector so long as the desired gene is capable of being transcribed and translated.

"Recombination" refers to the reassortment of sections of DNA or RNA sequences between two DNA or RNA molecules. "Homologous recombination" occurs between two DNA molecules which hybridize by virtue of homologous or complementary nucleotide sequences present in each DNA molecule.

The terms "stringent conditions" or "hybridization under stringent conditions" refers to conditions under which a probe will hybridize preferentially to its target subsequence, and to a lesser extent to, or not at all to, other sequences. "Stringent hybridization" and "stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments such as Southern and northern hybridizations are sequence dependent, and are different under different environmental parameters. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes part I chapter 2 Overview of principles of hybridization and the strategy of nucleic acid probe assays*, Elsevier, New York. Generally, highly stringent hybridization and wash conditions are selected to be about 5°C lower than the thermal

melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Very stringent conditions are selected to be equal to the T_m for a particular probe.

- 5 An example of stringent hybridization conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or northern blot is 50% formamide with 1 mg of heparin at 42°C, with the hybridization being carried out overnight. An example of highly stringent wash conditions is 0.15 M NaCl at 72°C for about 15 minutes. An example of stringent wash
- 10 conditions is a 0.2x SSC wash at 65°C for 15 minutes (see, Sambrook *et al.* (1989) *Molecular Cloning - A Laboratory Manual* (2nd ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, NY, for a description of SSC buffer). Often, a high stringency wash is preceded by a low stringency wash to remove background probe signal. An example medium stringency wash for a duplex of, *e.g.*, more than 100
- 15 nucleotides, is 1x SSC at 45°C for 15 minutes. An example low stringency wash for a duplex of, *e.g.*, more than 100 nucleotides, is 4-6x SSC at 40°C for 15 minutes. In general, a signal to noise ratio of 2x (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization.
- 20 Nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, *e.g.*, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code.

"Expression vectors" are defined herein as nucleic acid sequences that are direct the transcription of cloned copies of genes/cDNAs and/or the translation of their mRNAs in an appropriate host. Such vectors can be used to express genes or cDNAs in a variety of hosts such as bacteria, bluegreen algae, plant cells, insect cells and animal cells.

- 5 Expression vectors include, but are not limited to, cloning vectors, modified cloning vectors, specifically designed plasmids or viruses. Specifically designed vectors allow the shuttling of DNA between hosts, such as bacteria-yeast or bacteria-animal cells. An appropriately constructed expression vector preferably contains: an origin of replication for autonomous replication in a host cell, a selectable marker, optionally one or more
- 10 restriction enzyme sites, optionally one or more constitutive or inducible promoters. In preferred embodiments, an expression vector is a replicable DNA construct in which a DNA sequence encoding a one or more PKS and/or NRPS domains and/or modules is operably linked to suitable control sequences capable of effecting the expression of the products of these synthase and/or synthetases in a suitable host. Control sequences
- 15 include a transcriptional promoter, an optional operator sequence to control transcription and sequences which control the termination of transcription and translation, and so forth.

- The term "conservative substitution" is used in reference to proteins or peptides to reflect amino acid substitutions that do not substantially alter the activity (specificity or binding affinity) of the molecule. Typically conservative amino acid substitutions
- 20 involve substitution one amino acid for another amino acid with similar chemical properties (e.g. charge or hydrophobicity). The following six groups each contain amino acids that are typical conservative substitutions for one another: 1) Alanine (A), Serine (S), Threonine (T); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N),

Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); and 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

5 A "biological molecule that is a substrate for a polypeptide(s)" refers to a molecule that is chemically modified by one or more polypeptides encoded by the specified nucleotide sequence or amino acid. The "substrate" may be a native molecule that typically participates in the biosynthesis of a macrotetrolide, or can be any other molecule that can be similarly acted upon by the specified polypeptide(s) (e.g., a synthetic analog determined to be a substrate).

10 A "polymorphism" is a variation in the DNA sequence of some members of a species. A polymorphism is thus said to be "allelic," in that, due to the existence of the polymorphism, some members of a species may have the unmutated sequence (*i.e.* the original "allele") whereas other members may have a mutated sequence (*i.e.* the variant or mutant "allele"). In the simplest case, only one mutated sequence may exist, and the polymorphism is said to be diallelic. In the case of diallelic diploid organisms, three
15 genotypes are possible. They can be homozygous for one allele, homozygous for the other allele or heterozygous. In the case of diallelic haploid organisms, they can have one allele or the other, thus only two genotypes are possible. The occurrence of alternative mutations can give rise to triallelic, *etc.* polymorphisms. An allele may be referred to by the nucleotide(s) that comprise the mutation.

20 "Single nucleotide polymorphism" or "SNPs are defined by their characteristic attributes. A central attribute of such a polymorphism is that it contains a polymorphic site, "X," most preferably occupied by a single nucleotide, which is the site of the polymorphism's variation (Goelet and Knapp U.S. patent application Ser. No.

08/145,145). Methods of identifying SNPs are well known to those of skill in the art
(see, e.g., U.S. Patent 5,952,174).

Abbreviations used herein include LB, Luria-Bertani; nt, nucleotide; ORF, open
reading frame; PCR, polymerase chain reaction; PEG, polyethyleneglycol; PKS,

5 polyketide synthase; RBS, ribosomal binding site; Apr, apramycin; R, resistant; Th,
thiostrepton; WT, wild-type; and TS, temperature sensitive.

III. CLONING AND GENERAL CHARACTERIZATION OF THE NONACTIN BIOSYNTHESIS GENE CLUSTER

10 The macrotetrolides are a family of cyclic polyethers composed of four molecules
of enantiomeric nonactic acid or its homologues in a (+)(—)(+)(—)-ester linkage and
exhibit a broad spectrum of biological activities, ranging from antibacterial, antifungal,
antitumor, to immunosuppressive activity. Studies of macrotetrolide biosynthesis in
Streptomyces griseus have unambiguously established that they are of polyketide origin.
15 Of special interest is the proposal that macrotetrolide biosynthesis involves a pair of
enantiospecific polyketide pathways (Figure 1). This proposal is supported by (a) the
efficient incorporation of (6*R*,8*R*)- or (6*S*,8*S*)-2-methyl-6,8-dihydroxynon-2*E*-enoic acid
B and (±)-A into nonactin (C), (b) the isolation of both (+)- and (—)-A and their dimers
from *S. griseus* fermentation, and (c) the genetic and biochemical characterization of
20 NonS that catalyzes formation of (—)-A and its homologues from (—)-B and its
homologs. An intriguing corollary of this proposal is the synthesis of a set of
enantiomeric polyketides, such as B, presumably by a pair of polyketide synthases
(PKSs). To make B from the carboxylic acid precursors, the latter PKSs were originally
believed to have invoked (a) the rare use of succinate as an intact four-carbon fragment
25 (C3-C6) and (b) the derivation of a three-carbon unit (C7-C9) from two molecules of

acetate. As an initial step in testing this hypothesis, the present inventors initially cloned and heterologously expressed the macrotetrolide biosynthetic gene cluster, revealing a novel type of PKS for polyketide biosynthesis that consists of type II ketoacyl synthases (KSs) and ketoreductases (KRs) but lacks acyl carrier protein (ACP). Kwon et al., *J. Am. Chem. Soc.* 123, 3385 (2001); Smith et al., *Antimicrob Agents Chemother.* 2000, 44, 1809.

A 55-kb contiguous DNA region was cloned from *S. griseus* DSM40695 using the previously characterized *nonR* resistance gene as a probe. Inactivation of *nonS*, a gene adjacent to *nonR*, completely abolished macrotetrolide production, confirming that the cloned DNA encodes macrotetrolide biosynthesis. Nucleotide sequence analysis of a 30-kb fragment of the cloned region revealed 32 open reading frames. In addition to *nonR* and *nonS*, as well as other biosynthesis, resistance, and regulatory genes, the sequenced gene cluster consisted of 5 KSs and 4 KRs but apparently lacked an ACP (shown at Figure 2A). Smith et al., *Anti-microb Agents Chemother.* 2000, 44, 1809.

Traditionally, PKSs have been classified as one of three types. Type I PKSs are multifunctional enzymes that are organized into modules and type II PKSs are multienzyme complexes consisting of discrete, largely monofunctional proteins, both of which are found so far only in microorganisms. Despite the structural difference, type I and II PKSs share a high degree of amino acid sequence similarity and both types of PKSs use ACP to activate substrates and to channel the growing polyketide intermediates. Chalcone synthase and its homologues, also known as type III PKSs, are distributed mainly in plants and found very recently in microorganisms as well. They are different structurally and mechanistically from the former and are essentially condensing

enzymes that have no apparent amino acid sequence similarity to KSs of both type I and II PKSs. Type III PKSs lack ACP and act directly on the coenzyme A (CoA) ester of simple carboxylic acids. The 5 KSs identified within the macrotetrolide gene cluster and schematically shown in Fig. 2A are discrete proteins, three of which, *orf7*, *orf13* and *orf14*, are highly homologous to KS of type II PKSs and 2 of which, *orf9* and *orf10*, resemble KS of type II fatty acid synthases. The present inventors observed that if the sequenced region encompasses all the genes for macrotetrolide biosynthesis, the 5 KSs and 4 KRs must catalyze the assembly of (+)- and (—)-B from the carboxylic acid precursors in the absence of ACP, representing a novel type of PKS for polyketide biosynthesis.

To exclude the possibility that the ACP could reside outside the sequenced region, the inventors carried out a series of inactivation of orfs that flank the genes known to be essential for macrotetrolide biosynthesis, including the KSs and KRs as well as the *nonS* and *nonR* genes. Since genes encoding antibiotic production in microorganisms are often clustered, the inventors reasoned that sequential inactivation of orfs until the resultant mutants no longer exhibit macrotetrolide non-producing phenotype should allow the localization of the boundaries of the macrotetrolide biosynthetic gene cluster. Thus, inactivation of *orf2*, *orf3*, *orf27*, *orf28*, or *orf30* generated *S. griseus* mutants that show no difference in macrotetrolide production as compared to the wild type *S. griseus* strain. In contrast, macrotetrolide production is severely impaired or totally abolished in *S. griseus* mutants whose *orf4*, *orf6*, *orf7*, *orf9*, *orf10*, *orf11*, *orf12*, *orf13*, *orf14*, *orf15*, *orf16*, *orf19* (*nonS*), or *orf22* was inactivated. These results suggested that the boundaries

of the macrotetrolide biosynthetic gene cluster lie at *orf3* to *orf4* and *orf22* to *orf25*, respectively (Figure 2A).

To ascertain that the macrotetrolide PKS does not recruit an ACP residing at another position within the *S. griseus* chromosome and that the identified gene cluster is sufficient for macrotetrolide biosynthesis, the inventors demonstrated the production of macrotetrolides in a heterologous host. They chose *Streptomyces lividans* 1326 as a host because vector systems for the expression of PKS gene clusters in this organism have been well developed and used to produce various natural products. Thus, a 25-kb fragment harboring *orf4* to *orf26* was cloned under the *actI* promoter on an apramycin-resistant pSET152 derivative, yielding pBS2013. Introduction of pBS2013 into *S. lividans* 1326 afforded apramycin-resistant *S. lividans* 1326 (pBS2013) transformants that were cultured in AP medium. As shown in Figs. 2B-C, macrotetrolide production in *S. lividans* 1326 (pBS2013) was evident upon high performance liquid chromatography (HPLC) analysis of the resultant fermentation, as compared to that of the wild type *S. griseus* DSM40695 strain; these metabolites were absent from fermentation of the *S. lividans* 1326 (pSET152) negative control. The identity of the macrotetrolides produced were verified by electrospray mass spectroscopy analysis, yielding the characteristic molecular ions (m/e for $[M+Na]^+$) of 759.4241 for C ($C_{40}H_{64}O_{12}+Na$, calcd. 759.4297), 773.4230 for monactin (D) ($C_{41}H_{68}O_{12}+Na$, calcd. 773.4453), 787.4333 for dinactin (E) ($C_{42}H_{72}O_{12}+Na$, calcd. 787.4610), and 801.4930 for trinactin (F) ($C_{43}H_{76}O_{12}+Na$, calcd. 801.4767). It is noteworthy that *S. griseus* DSM40695 accumulates the macrotetrolides predominantly in mycelia (~ 40 mg/l). In contrast, *S. lividans* 1326 (pBS2013) preferentially secretes the macrotetrolides into medium (~ 10 mg/l) with only very small

quantity accumulated in mycelia (~ 0.5 mg/l). The latter observation was consistent with findings from the genome sequencing project of *Streptomyces coelicolor*, a very close cousin of *S. lividans*, that these organisms have an unusually high number of efflux pump proteins that efficiently export secondary metabolites out of the cells.

5 Macrotetrolide production in *S. lividans* 1326 (pBS2013) unambiguously demonstrated that the 23 orfs on pBS2013 are sufficient to support macrotetrolide biosynthesis from the carboxylic acid precursors, excluding the participation of other *S. griseus* gene products in macrotetrolide biosynthesis. It should be noted that this data fell short of excluding completely the possibility that the macrotetrolide PKS recruits an ACP from the hosts' fatty acid biosynthetic machinery. It is known that malonyl CoA-ACP transacylase from fatty acid biosynthetic pathway is an essential component of type II PKS. However, there is no evidence for crosstalk between fatty acid synthase ACP and PKS. Both type I or II PKSs known to date all include an ACP, either as a domain or a discrete protein, without which the PKSs are not functional. Therefore, the finding that 10 the macrotetrolide PKS consists of discrete KS and KR but lacks ACP uncovered an ACP-independent type II PKS for polyketide biosynthesis.

IV. IDENTIFICATION OF C-O BOND FORMATION BY A POLYKETIDE SYNTHASE

A. MATERIALS AND METHODS

20 Materials and methods related to the characterization of the *nonJKL* genes described in this section are presented below.

Standard Molecular Biology Techniques. Examples of such techniques and instructions sufficient to direct persons of skill through many cloning exercises described

herein are found in Berger and Kimmel (1989) *Guide to Molecular Cloning Techniques, Methods in Enzymology 152* Academic Press, Inc., San Diego, CA (Berger); Sambrook *et al.* (1989) *Molecular Cloning - A Laboratory Manual* (2nd ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, NY; Ausubel (1994) *Current*

5 *Protocols in Molecular Biology*, Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., U.S. Patent 5,017,478; and European Patent No. 0,246,864.

Gene inactivation and mutant complementation. Inactivation of *nonJKU* in *S. griseus* and *nonPQ* in *S. lividans* (pBS2013) by gene replacement, complementation of

10 the resultant mutants by either expression of the inactivated gene in trans or fermentation in the presence of exogenous (\pm)-3, and production of 1 by both the wild-type and recombinant *S. griseus* and *S. lividans* strains were carried out as reported previously for *nonS* (1) and the entire macrotetrolide biosynthesis gene cluster.

Construction of non gene expression cassettes pBS2014 to pBS2021 and

15 *biotransformation of* (\pm)-3 *into 1 in S. lividans 1326.* In general, the *actI* promoter was first combined with the desired *non* gene(s) in the *E. coli* vector, Lithmus 28 (NEB, Beverly, MA), and the resultant *actI/non* gene fragment was then transferred into *Streptomyces* shuttle vectors, pSET152 or pWHM3 (Fig. 9). The *actI* promoter was retrieved from pCK12 as a 3.2-kb of *NdeI-HindIII* fragment and ligated into the *EcoRV*

20 and *HindIII* sites of Litmus 28 to yield pBS2022. The *non* genes were first subcloned into an *E. coli* vector and then moved as *EcoRI-NsiI* fragment into the same sites of pBS2022. The resultant *actI/non* gene fragment was finally moved as a *XbaI-SpeI* fragment and cloned into the *XbaI* site of pSET152 or pWHM3 to afford the final

expression cassettes. As summarized in Fig. 6B, the *non* cluster (schematically illustrated in Fig. 6A) is dissected into two halves. Genes from the upstream half were cloned into pSET152, and genes from the downstream half were cloned into pWHM3, yielding the corresponding expression constructs pBS2014 to pBS2019. Expression of

5 *non* genes in both pSET152 and pWHM3-based cassettes were under the control of the *actI* promoter. For *nonL* expression, both the *actI* promoter and *snpA* promoter from pANT865 were used. Constructs with either promoter give similar biotransformation result.

These plasmids were introduced into *S. lividans* either pair wise or individually

10 and selected with apramycin (for pSET152-based plasmids) and/or thistrepton (for pWHM3-based plasmids). The resultant *S. lividans* recombinant strains were fermented (50 ml) in the presence of exogenously added (\pm)-3 (5 mg). Isolation and HPLC and ESI-MS analyses of 1 were carried out as previously described (1,2). The bulk of (\pm)-3 was recovered in the fermentation broth due to its poor cell permeability, reducing the

15 overall yield of 1 produced from these experiments.

Construction of the nonJ and nonK mutant expression cassettes pBS2020 and pBS2021. The NonK Cys161Gly and NonJ Cys169Gly mutants were generated by site-directed mutagenesis with the following pairs of primers 5'-

CGGTGAGCTGCGCGGCCTCCTCCTCGGTGC-3'/3'-

20 GCCACTCGACGCCCGCGGAGGAGGAGCCACG-5' (for *nonK*) and 5'-

GCGTGGCGGGCTCCGGCAATGTGGCGCTGCGG-3'/3'-

CGCACCGCCCGAGGCCGTTACACGCGACGCC-5' (for *nonJ*) (the Gly codons are underlined), respectively, using the QuickChange kit from Stratagene (La Jolla, CA)

according to the manufacture's instruction. The resultant *nonJK* mutants were cloned into pSET152 to yield pBS2020 and pBS2021, in which *nonJK* expression is under the control of the *actI* promoter. pBS2020 or pBS2021 was co-transformed with pBS2018 into *S. lividans*, and the resultant *S. lividans* recombinant strains were tested for

5 biotransformation of (\pm)-3 into 1 as described above.

Expression of nonL and characterization of NonL as a CoA ligase. The *nonL* gene was amplified from pBS2013 by PCR with forward primer of 5'-CGCCGGGGAGACCATATGATCGACGATGTGCTC-3' (the *NdeI* site is underlined) and reverse primer of 3'-GCATACTTGGTCCTTCTTAAGGCCCGCCGGTC-5' (the

10 *EcoRI* site is underlined) and cloned as a 1.7-kb *NdeI-EcoRI* fragment into the same sites of pET28a (Novagen, Madison, WI) to yield pBS2023. The expression of *nonL* in *E. coli* BL-21 (DE-3) (pBS2023) and purification of the resulting NonL protein by affinity chromatography on Ni-NTA resin were carried out under the standard conditions recommended by Novagen (Fig. 10). The incubation temperature was lowered to 15°C to

15 improve the solubility.

For *in vitro* assay of NonL as a CoA ligase, the complete assay solution (50 μ l) consisted of 2 mM ATP, 2 mM CoA, 2 mM (\pm)-3, 10 mM KCl, 2 mM DTT, 5 mM MgCl₂, 50 to 800 nM NonL in 10 mM Tris-HCl, pH 8.0. The reaction mixture was incubated at 30°C for 2 hr and analyzed by HPLC on an Alltima C₁₈ column (4.6 x 250

20 mm, 5 μ , 100Å, Alltech, Deerfield, IL). The column was first eluted in 20 mM NH₄Ac, pH 5.8, for 5 min, followed by linear gradient from 0% to 50% CH₃CN in 20 mM NH₄Ac, pH 5.8 for 10 min, and maintained at 50% CH₃CN in 20 mM NH₄Ac for an additional 10 min, with flow rate of 0.8 ml/min and detection at 260 nm.

B. BOND FORMATION BY POLYKETIDE SYNTHASES

As generally described in a previous section, PKSs have been classified according to their enzyme architecture, and three types of bacterial PKSs are known to date (Figs. 4A-C and 5A, B). Type I PKSs are multifunctional proteins that harbor sets of

- 5 noniteratively used distinct active sites, termed modules, each of which is responsible for the catalysis of one cycle of polyketide chain elongation. (Bacterial type I PKS that acts iteratively is also known but very rare and, therefore, is not included here.) A typical module consists minimally of an acyltransferase (AT) domain for extender unit selection and transfer, an ACP domain for extender unit loading, a ketoacyl synthase (KS) domain
- 10 for decarboxylative condensation between the aligned acyl thioesters to elongate the growing polyketide chain (Fig. 4A). Additional domains have also been identified for the modification of the initial β -keto group, such as a ketoreductase (KR) domain for a β -hydroxyl group, a dehydratase (DH) domain for an alkene moiety, an enoyl reductase (ER) domain for an alkane moiety, or an methyl transferase (macrotetrolide) domain for
- 15 the introduction of a methyl branch into the α -position of the growing polyketide chain, contributing additional structural diversities into polyketide biosynthesis. The modular structure of type I PKS and the co-linear relation between the order of domains and modules on each PKS and the functional groups in the resultant polyketide product have greatly facilitated the rational engineering of PKSs for combinatorial biosynthesis.
- 20 Numerous polyketides with predicted structural alterations have been produced by targeted domain or module substitution, deletion, addition, reposition, or by introduction of other mutations into PKSs, as well as by exploiting intermodular communication to

facilitate the transfer of biosynthetic intermediates between unnaturally linked PKS modules.

As best illustrated in Fig. 4B, Type II PKSs are multienzyme complexes that carry a single set of iteratively used activities and consist minimally of the KS α , KS β , and ACP subunits. (The KS β subunit lacks the conserved cysteine (Cys) residue essential for the KS activity and is also known as chain length factor or chain initiation factor.) Although self-malonylation of the ACP subunit has been observed *in vitro*, it is generally accepted that the minimal type II PKS requires the malonyl CoA:ACP transferase (MAT) from the host's fatty acid biosynthetic machinery to synthesize polyketide from malonyl CoA *in vivo*. MAT catalyzes the transfer of the malonyl group from malonyl CoA to the ACP subunit to form malonyl-ACP. To initiate polyketide biosynthesis, malonyl-ACP is decarboxylated by KS β to afford the starter unit acetyl-ACP that is subsequently transferred to KS α . Sequential decarboxylative condensations between the aligned acyl thioesters yield the linear poly- β -ketoacyl-ACP intermediate (Fig. 4B). Additional subunits have also been identified for the modification of the poly- β -ketoacyl-ACP intermediate, such as specific KRs, cyclases, and aromatases, furnishing the characteristic polycyclic aromatic products. Although the molecular logic of how type II PKSs control structural diversity in polyketide biosynthesis remains poorly understood, numerous novel aromatic polyketides have been produced by combinatorial biosynthesis via engineered type II PKSs. However, most of these experiments were governed mainly by intuition and serendipity. The so-called "design rules", while with enough predictive power for the rational production of novel aromatic polyketides by the appropriate mixing and matching of the subunits, were formulated exclusively from empirical

observations rather than mechanistic understanding of the type II PKS complex. Despite their structural difference, type I and type II PKSs share a high degree of amino acid sequence similarity, and both types of PKSs use ACP to activate the acyl CoA substrates and to channel the growing polyketide intermediates.

5 Now referring to Fig. 4C and Fig. 5B, Type III PKSs, also known as chalcone synthases-like PKSs, are distributed predominantly in plants and have been found in microorganisms only very recently. They are different from the former two types of PKSs both in structure and mechanism. Type III PKSs are essentially condensing enzymes. Although they possess a highly conserved Cys residue that is essential for the
10 PKS activity, the amino acid sequences of this Cys motif have no apparent similarity to that of the KSs of both type I and type II PKSs. Type III PKSs lack ACP and act directly on the acyl CoA substrates. They utilize the Cys active site iteratively by first selecting an acyl CoA starter unit and then catalyzing sequential decarboxylative condensation with malonyl CoA directly to synthesize a linear poly- β -ketoacyl CoA intermediate that
15 is subsequently cyclized into aromatic products (Fig. 4C). Although other subunits have been found for plant type III PKS, such as KR and methyl malonyl CoA-specific subunit, contributing additional structural diversity for plant polyketide biosynthesis, such subunit has not been identified for bacterial type III PKSs. Combinatorial biosynthesis with engineered type III PKSs were limited to plant enzymes so far. In contrast to type I and
20 type II PKSs, the relative simplicity of the type III PKSs has greatly facilitated their structural characterization. The crystal structures of two members of the type III PKS family, a chalcone synthase from *Medicago sativa* (alfalfa) and a 2-pyrone synthase from *Gerbera hybrida* (daisy), have been recently determined, revealing a common three-

dimensional fold, a set of conserved catalytic residues, similar CoA binding sites, but variable size of the active site cavity. The latter results imply that type III PKSs may control starter unit selectivity and limit the chain length of polyketide products by varying the size of the active site cavity. Changing the volume of the active site cavity through site-directed mutagenesis indeed has allowed alterations of both the choice of the starter unit and the final length of the resultant polyketide products, providing a structural basis for engineering type III PKSs for structural diversity

C. NOVEL C-O BOND FORMATION BY A POLYKETIDE SYNTHASE

In the present invention, the inventors have identified two KSs, NonJK, that catalyze C—O bond formation in nonactin (1) biosynthesis, as illustrated in Fig. 5C. In general, the NonJK KSs act on (\pm)-nonactyl CoA (2), catalyze tetramerization of (+)- and (-)-2 in a stereospecific (+)(-)(+)(-)-fashion into 1, and require the Cys residue conserved among all KSs for this reaction (Fig. 5C).

As described above, nonactin belongs to the macrotetrolide family of cyclic polyethers that exhibit a broad spectrum of biological activities, including antibacterial, antifungal, antitumor, and immunosuppressive activity. Referring to Fig. 5C, 1 is structurally composed of four molecules of (+)- and (-)-nonactic acid (3) in a (+)(-)(+)(-)-macrotetrolide linkage, an intriguing molecular topology not seen in other natural products. The biosynthesis of 1 has been studied by feeding experiments with various isotope-labeled precursors and by cloning and characterizing the biosynthesis gene cluster from *Streptomyces griseus* DSM40695 (previously described herein). These studies established that (\pm)-3 is synthesized by a PKS that lacks an ACP. The idea that (+)- and (-)-3 are intermediates in nonactin biosynthesis was supported by their efficient

incorporation into 1 and the isolation of both (+)- and (-)-3 as well as their dimers from *S. griseus* fermentation.

There are five KS genes within the cluster, *nonJKPQU*, which have deduced products that are highly homologous to KSs of the type II PKS (Figs. 2A and 6A).

5 Inactivation of any one of them either completely abolishes (*nonJKPQ*) or significantly impairs (*nonU*) 1 production, and expression of any one of them in trans in the corresponding mutants restores 1 production, confirming that all five KS genes are involved in 1 biosynthesis. Since KSs are only known to catalyze C—C bond forming steps in polyketide biosynthesis, the inventors reasoned that all KS mutants would retain
10 the enzymatic functions for the C—O bond-forming tetramerization steps involved in (±)-3 to 1 and, thereby, nonactin production could be restored to all KS mutants by fermenting them in the presence of exogenously added (±)-3. That was indeed the case for the *nonPQU* mutants but not for the *nonJK* mutants, suggesting that the NonJK KSs might play a role in the C—O bond-forming steps in nonactin biosynthesis.

15 To identify the minimal genes required for the C—O bond-forming tetramerization steps, the inventors combined plasmid-based expression of the *non* genes in *Streptomyces lividans* schematically illustrated in Figs. 6A-B with biotransformation of (±)-3 to 1 in the resultant recombinant strains. The inventors reasoned that the minimal genes required for the tetramerization step could be identified by successive
20 deletion of genes included in the expression cassettes. Thus, *S. lividans* strains harboring various *non* gene expression cassettes were fermented in the presence of exogenously added (±)-3. The production of 1 was monitored by high performance liquid chromatography (HPLC) and confirmed by electrospray ionization-mass spectrometry

(ESI-MS) analysis, yielding a characteristic $(M+Na)^+$ ion at $m/z = 759.4$, consistent with the molecular formula of $C_{40}H_{64}O_{12}$ for 1. As summarized in Fig. 6B and 6C, all recombinant strains expressing *nonJKL* genes are effective in biotransformation of (\pm) -3 into 1 (entries II, VI, VII, and VIII) and deletion of *nonJK*, or *nonL* from the expression cassettes completely abolishes their biotransformation ability (entries III, IV, and V), suggesting *nonJKL* are essential for this activity. Finally, it was established that *nonJKL* are sufficient for the tetramerization steps by expressing only the *nonJKL* genes; the resultant recombinant strain is as effective in biotransformation of (\pm) -3 into 1 (entry VIII) as those strains harboring other *non* genes in addition to *nonJKL*. The analysis demonstrates that a substantial portion of the nonactin gene cluster may be absent while the biotransformation of (\pm) -3 to 1 is still carried out. The nucleotide sequences for the *nonJ*, *nonK* and *nonL* open reading frames are provided herein as SEQ ID NOS. 2, 4 and 6, respectively. The translated polypeptides for nonJ, nonK and nonL are likewise provided as SEQ ID NOS. 3, 5 and 7. SEQ ID NO. 1 is approximately 15 Kb of genomic DNA from the nonactin biosynthesis gene cluster which contains the open reading frames set forth in SEQ ID NOS. 2, 4 and 6. SEQ ID NO. 1 has been deposited with GenBank under Accession No. AF074603. The entire nonactin biosynthesis gene cluster has been deposited with GenBank under Accession Nos. AF263011 and AF263012, along with the deposit of Accession No. AF074603. Cloning of the nonactin biosynthesis gene cluster has previously been described in Smith, W.; Xiang, L.; Shen, B. *Antimicrob. Agents Chemother.* 2000, 44, 1809-1817 and Kwon, H. J., Smith, W.C.; Longkuan X., and Shen, B., *J. Am. Chem. Soc.* 2001, 123, 3385-3386 which are incorporated herein by reference in their entirety for all purposes.

On the basis of its high sequence homology to a family of CoA ligases, the inventors reasoned that NonL, rather than directly contributing to the C—O bond-forming steps, instead activates (\pm)-3 into CoA esters (\pm)-2 that are then tetramerized into 1 by the NonJK KSs. To validate this hypothesis, the inventors expressed *nonL* in *E. coli*, purified NonL as a His₆-tagged fusion protein (Fig. 10), and characterized NonL as a CoA ligase. The inventors monitored the *in vitro* assay of NonL by HPLC analysis: NonL catalyzes the conversion of (\pm)-3 to (\pm)-2, requiring CoA and ATP as co-substrates. Under the condition examined, 80% of (\pm)-3 was converted to (\pm)-2 in 2 hrs, confirming that NonL recognizes both (+)- and (-)-3 as substrates (Fig. 7). The identity of (\pm)-2 was verified by ESI-MS, yielding a characteristic (M-2H)²⁺ ion at m/z = 474.7, consistent with the molecular formula of C₃₁H₅₂N₇O₁₉P₃S.

Since NonJKL are sufficient to biotransform (\pm)-3 to 1 (Fig. 6B and 6C, entry VIII), the characterization of NonL as a CoA ligase, catalyzing the conversion of both (+)- and (-)-3 into their CoA esters, indicates that NonJK are responsible for the C—O bond-forming tetramerization steps, acting directly on the CoA ester substrates of (+)- and (-)-2. KSs known to date catalyze only C—C bond formation, although the CoA substrates can be utilized either directly (by type III PKS) or indirectly via acyl-ACPs (by both type I and -II PKS). Thus, as depicted in Fig. 5C, NonJK, in a mechanistic analogy to type III PKS, it was suggested that NonJK first catalyze the transfer of the nonactyl group from (\pm)-2 to the Cys residue of NonJK to form the nonactyl-S-KS species (Fig. 5C, step a). Unlike KSs that catalyze decarboxylative condensation between the carbon anion nucleophile and acyl-S-KS to form the C—C bond (Fig. 5B), the NonJK KSs catalyze the condensation between the oxygen nucleophile of the -OH group of 2 and

nonactyl-S-KS to form the C—O bond, yielding the dimers in the form of CoA esters (Fig. 5C, steps b and c). The latter have been isolated as free acids from *S. griseus* fermentation, supporting the proposed pathway. Iterations of steps a and b or steps a and c eventually lead to the KS-bound linear tetramers that undergo intramolecular

5 condensation between the —OH group of the distal nonactyl unit and the acyl-S-KS carbonyl group to afford 1 (Fig. 5C, step d).

To gain insight into the mechanism of the condensation reaction, the inventors further compared the NonJK sequences with those of other KSs of both fatty acid synthases (FASs) and PKSs. All previously characterized KSs contain a Cys-His-His

10 (for type I and —II FAS and PKS) or Cys-His-Asn (for type III PKS) catalytic triad (Fig. 8). The His-His or His-Asn residues are essential for malonyl-ACP or malonyl CoA decarboxylation to generate the corresponding carbon anion, and the Cys residue catalyzes condensation between the resultant carbon anion and acyl-S-KS to form the C—C bond (Fig. 5A and 5B). Strikingly, NonJK are characterized with a mutated

15 catalytic triad—Cys-Tyr-His for NonJ or Cys-Gly-His for NonK, suggesting that NonJK lack the decarboxylation activity (Fig. 8). This is consistent with the proposal that NonJK catalyze C—O bond formation by using the —OH as the nucleophile directly.

Finally, to confirm that Cys plays a catalytic role in the C—O bond-forming step, the inventors replaced the conserved Cys residue in NonJ or NonK, respectively, with

20 Gly by site-directed mutagenesis. The resultant mutants completely lose their ability to biotransform (\pm)-3 into 1 (Fig. 6B and 6C, entries IX and X). Thus, the NonJK KSs catalyze the C—O bond-forming step in nonactin biosynthesis, acting directly on (\pm)-2

and using the same active site residue Cys that is used in KS catalysis of C—C bond formation.

In summary, nucleotide sequence analysis of a 30-kb of the macrotetrolide biosynthetic gene cluster revealed thirty two open reading frames (orfs) composed of five discrete ketosynthases (KSs) and four ketoreductases (KRs) but lacks an ACP. The GenBank accession numbers for the macrotetrolide biosynthetic gene cluster are AF263011, AF263012, and AF074603A 25-kb DNA region encompassing 24 orfs conferred macrotetrolide production to *S. lividans* 1326, which substantiated that macrotetrolide-PKS is a ACP-independent type II PKS composed of five KSs and four KRs. A gene of *nonY* was recently identified just upstream to *nonX*. The insert of pBS2013 covers for *nonY*. Subsequently, two of five KS genes, *nonJK* were characterized to encode for the conversion of (\pm)-6-CoA to 1. (See Fig. 11). NonJK was identified as novel KSs catalyzing C-O bond formations between (\pm)-6-CoA to yield 1 and NonL as an enzyme converting (\pm)-6 into (\pm)-6-CoA. The chemistry of NonJK-catalysis was further defined by the inventors revealing stereoselective C-O bond formations performed in sequential order of NonJ-NonK, as described below.

Close examination of *nonK* predicted two plausible translational starting sites. Frame plot analysis predicted *nonK* encodes a 462-aa protein (VSKEH-NonK) starting from a GTG-starting codon and translationally-coupled (4-bp) with *nonN*, the upstream KR-gene. At 111-nucleotide downstream from this GTG-starting codon, the other starting codon of ATG was found and the product deduced from this ATG-codon (MGFCL-NonK) showed head-to-tail homology to the known KSs from type II polyketide synthase (PKS) and fatty acid synthase (FAS). Through polymerase chain reaction (PCR), *nonK*

was amplified, from both starting sites, along with *nonJ* that is translationally-coupled with *nonK* (11-bp) and subcloned into Lithmus28 to be combined with *nonL* generating pBS2043 and pBS2044, respectively (see supporting information for plasmid constructions). Firstly, the inserts (NonK, NonJ, and NonL) were rescued as *NdeI-BglII* fragments and ligated with *NdeI-BamHI* digested pIJ4123 to be located downstream to the thiostrepton inducible promoter (*tipAp*), ribosomal binding site/translational starting codon, and 6×histidine-tag. The ligation mixtures (pBS043/pIJ4123 and pBS2044/pIJ4123, respectively for VSKEH-NonK and MGFCL-NonK) were used to transform *S. lividans* TK-24 and the transformant was selected by kanamycin resistance and later by loss of red pigment production, due to loss of *redD*, on the regeneration agar. Through plasmid isolation and restriction digestion, *S. lividans* (pBS2045) was identified to harbor pBS2044/pIJ4123 (two out of twelve tested in the liquid cultures, two of five not showing red pigment production even after the thiostrepton-supplementation) but no transformant was found with pBS2043/pIJ4123. Tests for the *in vivo* conversion-activity of *S. lividans* (pBS2045) clearly indicated that *S. lividans* (pBS2045) could not convert (±)-6 into 1. The inventors also subcloned the inserts of pBS2043 and pBS2044 into pET28a to generate pBS2046 and pBS2047. *E. coli* BL21 (pBS2046) accumulated low level of 1 when supplemented by (±)-6 after IPTG (isopropyl-β-D-thiogalactopyranoside)-induction. Consistent with the result observed with *S. lividans* (pBS2045), no conversion activity was observed from *E. coli* BL21 (pBS2047). SDS (sodium dodecyl sulfate)-PAGE (polyacrylamide gel electrophoresis) analysis revealed that both versions of NonKs were well expressed but almost insoluble and trial to resolve 6×histidine-tagged

NonK or any relevant activities through the nickel affinity chromatography have been unsuccessful.

At this point, the inventors proceeded with cell-free extract (CFE)-based activity assay. CFE of *S. lividans* (pBS2045) and *E. coli* BL21 (pBS2046) were incubated with ATP, CoA, and (\pm)-6 (see supporting information for CFE-preparation and chemical analysis). *S. lividans* (pBS2045)-CFE showed no conversion to 1, as expected from *in vivo* results, but instead yielded comparable levels of 6-CoA, 8-CoA, trimeric 6-CoA and tetrameric 6-CoA as evidenced by MS analysis that revealed, respectively, molecular ion (m/e for [M-H]⁺) of 950.207 (C₃₁H₅₁N₇O₁₉SP₃, calcd 950.217), 1134.336 (C₄₁H₆₇N₇O₂₂SP₃, calcd 1134.327), 1318.430 (C₅₁H₈₃N₇O₂₅SP₃, calcd 1318.437), and 1502.564 (C₆₁H₉₉N₇O₂₈SP₃, calcd 1502.547) (Figure 12A-III). Production of these series of compounds was dependent on 6 (Figure 12A-II) and ATP. However, *E. coli* BL21 (pBS2046)-CFE also failed to produce 1 but yielded 6-CoA and a trace level of 8-CoA (Figure 12A-IV). The results indicated that NonL was well expressed in *E. coli* BL21 (pBS2046) but an enzyme for 8-CoA formation was limitedly expressed, which might be the reason for absence of the *in vitro* conversion activity to 1. The result from *S. lividans* (pBS2045)-CFE (Figure 12A-III) suggested that oligomerization of 6-CoA occurs step-by-step with the intermediacy of trimeric 6-CoA and MGFL-NonK can support oligomerization but not intramolecular cyclization that necessitate full-length of NonK of VSKEH-NonK. From *S. lividans* (pBS2045)-CFE, neither relevant activity nor protein(s) could be resolved with the nickel affinity chromatography. However, surprisingly, the flow-through fraction from the nickel affinity chromatography showed efficient conversion of 6-CoA to 8-CoA in addition to NonL activity (Figure 12B). It was assumed

that removal of association of MGFCL-NonK from NonJ, somehow during nickel affinity chromatography, refined NonJ for its catalysis, the dimerization. This result made the inventors propose that NonJ independently catalyzes dimerization of 6-CoA. The absence of *in vitro* conversion ability in *E. coli* BL21 (pBS2046) was interpreted as due to lack of NonJ whose translation is dependent on its native sequence of *Streptomyces* ribosomal binding site.

To further evaluate this mechanism, it was decided to establish an *in vitro* conversion system that can readily convert (\pm)-6 into 1. For this purpose, ActIIORF4-activated *actII/actIII* promoter system was adopted to achieve concerted expression of the components, NonKJ. The NonKJL-expression plasmids were therefore constructed on pWHM467 and introduced into *S. lividans* for *in vivo* and *in vitro* assays; pBS2048 and pBS2049 were prepared for NonJ, NonL, and respectively for VSKEH-NonK and MGFCL-NonK for *in vivo* conversion experiments. While *S. lividans* (pBS2048) efficiently converted (\pm)-6 into 1, *S. lividans* (pBS2049) could not, defining that NonK is, at least as the catalyst supporting conversion of (\pm)-6 to 1, 462-aa protein with unusual N-terminal extension. Consistent with the *in vivo* feeding experiment, 1 was evident in the ethyl acetate extract of *S. lividans* (pBS2048)-CFE but not found from *S. lividans* (pBS2049)-CFE (Figure 13AB). With *S. lividans* (pBS2048)-CFE, a 10 % of yield was achieved for conversion of (\pm)-6 to 1 within 1h. From both of *S. lividans* (pBS2047)-CFE and *S. lividans*(pBS2048)-CFE, 6-CoA and 8-CoA were found and a large excess of 6-CoA was accumulated compared to 8-CoA (Figure 13C). The difference between results from *S. lividans* (pBS2045) (Figure 12A-III) and *S. lividans* (pBS2048) (Figure 13C-IV) was attributed to difference in the engineered transcriptional systems and whereby, the

engineered actinorhodin promoters were viewed as rendering the concerted transcription of *nonKJ*. To further substantiate the roles of 6-CoA and 8-CoA in the nonactin biosynthesis, each compound was purified and incubated with *S. lividans*(pBS2048)-CFE and *S. lividans*(pBS2049)-CFE. 6-CoA and 8-CoA were collected from HPLC, freeze-dried and roughly quantified by using the known extinction coefficient of free coenzyme A (16,800 at 259.5 nm). When 6-CoA was incubated with *S. lividans* (pBS2048)-CFE, a small portion of 6-CoA could be seen intact, without any detectable level of 8-CoA, with accumulation of 1. Contrarily, in *S. lividans* (pBS2049)-CFE supplemented with 6-CoA, the comparable levels of 6-CoA and 8-CoA were found but 1 was not found. *S. lividans* (pBS2048)-CFE achieved quantitative conversion ($\geq 50\%$) of either of 6-CoA (60 μM) or 8-CoA (40 μM) into 1 within 1h. The efficient conversion of 8-CoA into 1 eventually disproved the intermediacy of trimeric 6-CoA but indicated 8-CoA is dimerized and subsequently cyclized to 1, by VSKEH-NonK.

To exclude any possibility that truncated NonK (MGFCL-NonK) or NonL charperoned NonJ-catalysis, *nonJ* alone was expressed in *S. lividans* and the resulting transformant was tested for the ability to convert 6-CoA to 8-CoA. Through PCR-amplification, *nonK* and *nonJ* were separately amplified and subcloned into pWHM3 downstream to *ermE* promoter generating pBS2053 and pBS2054, respectively. NonL was incubated with ATP, CoA, and 6 and then boiled for enzyme inactivation. When this NonL-reaction mixture was used, production of 8-CoA was evident with *S. lividans* (pBS2054)-CFE but not with *S. lividans* (pBS2053)-CFE (Figure 14A). It should be noted that two molecules of 6-CoA are consumed to form one molecule of 8-CoA, assuring that NonJ is enough to convert 6-CoA to 8-CoA and NonK and NonL has no

relevant role in this conversion process. The production of 1 can be achieved by incubating 8-CoA with *S. lividans*(pBS2053)-CFE, which further assured the sequential catalysis of NonJ and NonK. However, the activity of *S. lividans*(pBS2053)-CFE was significantly compromised compared to *S. lividans* (pBS2048), which provided the possibility of NonJ-chaperoning on NonK catalysis. The CFE of *E. coli* BL21 (pBS2047) showed efficient NonK activity comparable to that of *S. lividans* (pBS2048) although it possesses just a trace level of NonJ activity (Figure 12A-IV). As a purely genetic concept, the inventors could conceive that the integrity of *nonK* required *nonJ* sequence on its downstream for its proper expression to NonK, for which there is apparently no clear explanation in the biochemical level, however. Alternatively, and more plausibly, a catalytic role of NonJ in NonK-catalysis can be envisioned; that is, NonJ takes up again 8 from 8-CoA and directs transfer of 8 onto the nucleophile of NonK similar to what has been proposed in aflatoxin biosynthesis.

The inventors then addressed the question of what is NonJ's stereospecificity toward the substrates, (-)-6-CoA and (+)-6-CoA. (+)-6-CoA or (-)-6-CoA was readily generated from cognate enantiomeric 6 by NonL with the same efficiency in the overall turn-over, which verified that NonL operates in the nonstereospecific manner as aforementioned. The deduced mechanism for NonJ catalysis is as follows: one enantiomeric 6-CoA donates a 6-moiety to the cysteine nucleophile of NonJ and a terminal hydroxyl group from the other enantiomeric 6-CoA replaces the active site nucleophile to generate 8-CoA. That is, the thioesteric linkage of one enantiomeric 6-CoA acts as a leaving group and the CoA-moiety of the other enantiomeric 6-CoA is maintained in 8-CoA(Figure 11). Therefore, to differentiate fates of (-)-6-CoA and (+)-6-

CoA, the investors labeled CoA-moieties of (-)-6-CoA and (+)-6-CoA differentially, by employing enantiopure 6 in a NonL reaction. To achieve this, the inventors reviewed the known CoA biosynthetic pathway and found that the ultimate intermediate to CoA is 3'-dephosphiu-CoA (dp CoA), which is commercially available. The formation of 6-dp-

5 CoA and 6-dp-CoA was carried out by employing purified NonL and the CFE containing NonJ. It was then demonstrated that dp-CoA can be converted into 6-dp-CoA by NonL with efficiency comparable to that with CoA and *S. lividans* (pBS2054)-CFE also converted 6-dp-CoA into 8-dp-CoA whose identity was verified by detecting molecular ion (*m/e* for [M-H]⁺) of 1054.368 (C₄₁H₆₆N₇O₁₉SP₂, calcd 1054.361). With dp-CoA to

10 replace CoA in NonL and NonJ catalysis, (-)-6-dp-CoA, (+)-6-dp-CoA, (-)-6-CoA, and (+)-6-CoA were separately prepared and incubated (100 μM each) with *S. lividans* (pBS2054)-CFE. HPLC analysis revealed that incubation of (-)-dp-6-CoA with (+)-6-CoA resulted in 8-CoA, whereas (-)-6-CoA with (+)-dp-6-CoA resulted in 8-dp-CoA (Figure 13B). A notable feature found in the (-)-dp-6-CoA/(+)-6-CoA-incubation was

15 that (-)-dp-6-CoA was consumed remarkably higher than (+)-6-CoA, which was attributed to hydrolysis of (-)-dp-6-CoA activated by the presence of (+)-6-CoA but uncoupled with formation of 8-CoA. The NonJ-catalysis was defined as stereoselective, in which (-)-6 was loaded onto the active site nucleophile that is replaced by terminal hydroxyl group of (+)-6-CoA to generated (-)-6-(+)-6-CoA. This experiment showed that

20 (+)-6-(+)-6-CoA also can be generated by NonJ with four fold less efficiency of that of (-)-6-(+)-6-CoA, whereas no trace of (-)-6-(-)-6-CoA can be generated. The (+)-6-(+)-6-CoA apparently can not lead to 1 as evident from the established structure of 1 and other

CFE-conversion experiments. It seems therefore evident that NonK discriminates (+)-6-(+)-6-CoA in the oligotetramerization/cyclization process.

NonJK are unique KSs that catalyze C-O bond formation. By exploiting the *in vitro* assay system, and CoA-derivatives, 6-CoA and 8-CoA were defined as direct
 5 substrates for NonJ and NonK catalysis, respectively. This feature was originally suggested by the absence of ACP in the macrotetrolide gene cluster and here verified by efficient conversion of 6-CoA into 8-CoA which then lead to 1 *in vitro*. So far, KSs directly accepting directly acyl-CoAs are known to have a primary structure clearly distinguishable from ACP-dependent ones. NonJK definitely belong to ACP-dependent
 10 KSs in the primary sequence analysis but here they are established as ACP-independent ones, demonstrating that catalytic mechanism (ACP-dependent or -independent) is surprisingly not related to the observed primary sequence features. Furthermore, the nucleophiles in condensations of NonJ and NonK are complex acyl group (6 and 8) other than malonyl group that is the general nucleophile in type II PKS and FAS and type III
 15 PKS (in the case of type III PKS, incorporation of methylmalonyl group occurs exceptionally). Furthermore, the inventors have demonstrated that NonJ and NonK work in a processive manner instead of in an iterative manner as that known for their homologues in type II and type III PKSs.

Materials and methods related to the characterization of *nonJKL* genes described
 20 in the foregoing section are presented below:

Plasmid preparation: The relevant *nonKJ* sequences were amplified from pBS2019 by Vent polymerase (NEB, Beverly, MA) with forward primer 5'-

TGGACGCGGGGGCCATATGAGCAAGAG-3' (the *NdeI* site is underlined) for VSKEH-NonK or 5'-CGCGCTGGTCACCCATATGGGGTCTGC-3' (the *NdeI* site is underlined) for MGFCL-NonK and reverse primer of 5'-GCCGCGTCGCCATGCATTGAACGTGGGT-3' (the *NsiI* site is underlined) and

5 cloned as a 2.7-kb or 2.6-kb *NdeI-NsiI* fragment into the same of pGEM-5zf (Promega, Madison, WI) generating pBS2041 and pBS2042. The *nonLS* was subcloned from pBS2003 as a 4.2-kb *KpnI-HincII* fragment into the same of pUC18. From the resulting plasmid, the insert was rescued as *EcoRI-HindIII* fragment and subcloned into the same of Lithmus 28 generating pHJK-3-45D. The inserts of pBS2041 and pBS2042 were

10 rescued as *SpeI-NsiI* fragment and subcloned into same of pHJK-3-45D to generate pBS2043 and pBS2044, respectively. The insert of pBS2044 was rescued as *NdeI-BglII* fragment and ligated into *NdeI-BamHI* sites of pIJ4123 to generate pBS2045.

The *NdeI-BglII* fragments including *nonKJL* were rescued from pBS2043 and pBS2044, respectively, and subcloned into *NdeI* and *BamHI* sites of pET28a (Novagen,

15 Madison, WI) generating pBS2046 and pBS2047. The inserts of pBS2046 and pBS2047 were recovered as *XbaI-NsiI* fragments and subcloned into pHJK-3-97A, a pWHM467 derivative to which the unique *BglII* site was engineered downstream to *PacI* site to generate pHJK-4-07E and pHJK-4-07F, respectively. The *nonDL* sequence was rescued as *PacI-BglII* fragment from pBS2016 and subcloned into same of pHJK-4-07E and

20 pHJK-4-07F, respectively, to generate pBS2048 (pHJK-4-10A) and pBS2049 (pHJK-4-10B).

The *nonK* sequence was amplified from pBS2019 by Vent polymerase with forward primer of 5'-

CCTCAGGCCCCATGGTCTAGAGCACCATCCTGCGGCGCCTG-3' (the *Nco*I and *Xba*I sites are underlined) and reverse primer of 5'-

GCAGAGGCAGATCTGCAGACATCGCCACCTCCCA-3' (the *Bgl*II site is

underlined). The *nonJ* was amplified from pBS2019 by Vent polymerase with forward

5 primer of 5'-GACCCCGTCCATGGTCTAGACATTCGACCCGGTCCCCGGC-3' (the

*Nco*I and *Xba*I sites are underlined) and reverse primer of 5'-

TGTAACGTAGATCTTGGCAAGTCGCCGCCTTCGT-3' (the *Bgl*II site is underlined).

The PCR products were purified as *Nco*I-*Bgl*II fragment and subcloned into the same of

pQE60 generating pBS2050 (pHJK-4-16C) (*nonK*) and pBS2051 (pHJK-4-16D) (*nonJ*).

10 The *ermE* promoter DNA (EMBL accession number X02392 for *ermE*) was recovered as a 300-bp *Kpn*I-*Bam*HI fragment from pIJ2925 and subcloned into pUC18 generating pBS2052 (pHJK-4-15A). The insert of pBS2050 (*nonK*) or pBS2051 (*nonJ*) was recovered as *Xba*I-*Hind*III fragments and the *ermE* promoter was rescued as an *Eco*RI-*Xba*I fragment from pBS2052, both of which were ligated into *Eco*RI and *Hind*III sites of pWHM3 to generate pBS2053 and pBS2054, respectively for *nonK* and *nonJ* expression.

20 *Nickel affinity chromatography*: SDS (sodium dodecyl sulfate)-PAGE (polyacrylamide gel electrophoresis) analysis of *S. lividans* (pBS2045)-cell extract showed, in insoluble fraction, a strong protein band around 45-kDa that was absent from controls such as that of *S. lividans*(pIJ4123) with thiostrepton-supplementation or *S. lividans* (pBS2045) before thiostrepton-supplementation. It was conceived that the highly expressed insoluble protein was MGFCL-NonK based on the size of 6×histidine-tagged-MGFCL-NonK as seen in *E. coli*BL21 (pBS2047). In the soluble fraction was

observed protein-enrichment, unresolved, around 45 to 50-kDa. Neither relevant activities (such as formation of 6-CoA and its oligomerization) nor protein(s) around 45 to 50-kDa was retained on the nickel affinity chromatography even in the modified condition.

5 The inserts of pBS2043 and pBS2044 were also subcloned into pET28a to generated pBS2046 and pBS2047. The expression experiment in *E.coli* showed that both versions of NonK were well expressed, without any relevant protein bands for NonJ or NonL, but largely insoluble and somehow, not compatible to the nickel affinity chromatography. It could not be determined whether the failure was simply due to
10 inability of 6×histidine-tagged-NonKs to interact with nickel or their intrinsic instability during the chromatographic procedure.

Cell-free extract (CFE) preparation: S. lividans (pBS2045) were grown on Yeast Extract-Malt Extract (YEME, 34 % sucrose) media supplemented with kanamycin (50µg/ml), at 30°C and 300rpm for 1 d and added with thiostrepton (5µg/ml). The
15 thiostrepton-treated culture was maintained in the same condition for 1 d and then harvested by centrifugation (13,600g, 20 min, 4°C). Cell were lysed in 30 ml of 100 mM sodium phosphate buffer (pH, 8.0)-1 mM EDTA-50 mM KCl-10 % glycerol with lysozyme (2mg/ml) for 30 min at 4°C, sonicated for thirty 10-sec bursts at 300W with a 30-sec cooling period between each burst. Centrifugation (27,500g, 20 min, 4°C) yield a
20 clear supernant that was further brought to 70 % saturation of (NH₄)₂SO₄ and centrifuged as above to the collect pellet. The pellet was dissolved in 20 mM Tris-Cl (pH, 8.0)-10 mM MgCl₂-2 mM DTT-10 mM KCl-10 % glycerol and desalted on Sephadex G-25 column to yield the cell-free preparation (7 to 10mg/ml proteins). The CFEs of *S.*

lividans(pBS2047) and *S. lividans* (pBS2048) were prepared at the same procedure except that they were grown on 25ml-AP media, supplemented with thiostrepton (5µg/ml), at 30°C and 300rpm for 3 ds. Similarly, the CFEs of *S. lividans* (pBS2053) and *S. lividans* (pBS2054) were prepared by using mycelium from YEME (34 % sucrose) cultures.

Conditions for CFE-assay and chemical analysis: The assay solution consisted of 1 to 2 mM each of (\pm)-6, CoA, and ATP and 0.8 mg/ml of protein in 10 mM Tris-Cl (pH, 8.0)-10 mM KCl-2 mM DTT-5 mM MgCl₂, and incubated at 30°C for 1 h to 3hs. The incubation was terminated by ethyl acetate extraction. The organic layer was analyzed for 1 by Thin-Layer Chromatography (TLC), Evaporative Light-Scattering Detector (ELSD)-High Performance Liquid Chromatography (HPLC), and positive Matrix-Assisted Laser Desorption Ionization (MALDI)-Mass Spectrometry (MS) analysis as previously described.^{5d} The aqueous layer was analyzed by HPLC with UV-detection, from which relevant peaks were collected and subjected to negative MALDI-MS analysis. HPLC with UV-detection was performed on an Alltima C18 column (4.6 x 250 mm, 5 µ, 100 Å, Alltech, Deerfield, IL). The column was first eluted in 5 % CH₃CN in 20 mM NH₄Ac, pH 5.8 (buffer A) for 7 min, followed by linear gradients from 100 % buffer A to 50:50 of (buffer A:CH₃CN) for 8 min, from 50:50 of (buffer A:CH₃CN) to 10:90 of (buffer A:CH₃CN) and then maintained for additional 10 min, except for the result in Figure 12A for which the elution was maintained after reach to 50 % CH₃CN, with flow rate of 1.0 ml/min at 260 nm. MALDI-MS analysis was performed on a HiResMALDI FT-Mass Spectrometer with 7 tesla superconducting magnet (IonSpec Corp., Irvine, CA) at the

Analytical Instrumentation Facility of School of Pharmacy, University of Wisconsin, Madison.

Non-stereoselective features in NonJ catalysis: The eluent from 8-CoA and 8-dp-CoA was collected from HPLC and analyzed by MS analysis, which confirmed their identities. The finding from mass spectra is that 8-dp-CoA molecular ion was found as a 25 % *rel. int.* of 8-CoA in (-)-dp-6-CoA/(+)-6-CoA and vice versa, 8-CoA found as a 5 % *rel. int.* of 8-dp-CoA in (+)-dp-6-CoA/(-)-6-CoA. Roughly, 8-dp-CoA is two times sensitive as 8-CoA in the MS analysis, which projected that about 10 % of 8-CoA has the stereochemistry [(+)(-)-CoA] opposite to that of major product, (-)-6-(+)-6-CoA. However, it should be also questioned whether the intensities on mass spectra do linearly correlate to the actual concentration of the specific molecule.

NonJ also can form (+)-6-(+)-6-CoA. NonJ accumulated significant level of 8-CoA or 8-dp-CoA from (+)-6 incubation with ATP and CoA/dp-CoA in *S. lividans* (pBS2045) and (+)-6-CoA and/or (+)-6-dp-CoA in *S. lividans*(pBS2048). Although the kinetic values for (-)-6-(+)-6-CoA and (+)-6-(+)-6-CoA were not determined, the formation of (-)-6-(+)-6-CoA or (-)-6-(+)-6-dp-CoA was observed to be two to four fold higher than that of (+)-6-(+)-6-CoA or (+)-6-(+)-6-dp-CoA in the overall turnover.

D. CREATION OF CHEMICAL DIVERSITY BY UTILIZING NonJK IN DIRECTED BIOSYNTHESIS.

Elements (*e.g.*, open reading frames) of the nonactin biosynthetic gene cluster, namely, *nonJ*, *nonK*, *nonJK* or *nonJKL* or variants thereof, can be used in a wide variety of "directed" biosynthetic processes (where the process is designed to modify and/or synthesize one or more particular preselected target molecule(s). Variants may include complete or partial open reading frames such as those encompassing active catalytic sites,

and the like. In particular, the *nonJKL* open reading frames can be used to synthesize a macrotetrolide or a macrotetrolide analogue. The combination of *nonJKL* open reading frames can be used to direct the synthesis of various macrotetrolides from starting products exemplified by 3 in Fig. 5C.

5 In addition, polypeptides from *nonJ* or *nonK* open reading frames, taken individually, can be used to perform chemical modifications on particular substrates and/or to synthesize various metabolites. Thus, for example, polypeptides of *nonJ* or *nonK* can be used to catalyze a C-O bond in a biological molecule appropriate as a substrate for the particular polypeptide whereby a C-O bond is catalyzed in
10 intermolecular or intramolecular fashion within the biological molecule or between the biological molecule and a second biological molecule. One of skill in the art, utilizing the information provided here, can perform literally countless chemical modifications and/or syntheses using either "native" macrotetrolide biosynthesis metabolites as the substrate molecule, or other molecules capable of acting as substrates for the particular
15 polypeptide. Other substrates can be identified by routine screening. Methods of screening enzymes for specific activity against particular substrates are well known to those of skill in the art.

 The biosyntheses can be performed *in vivo*, e.g. by providing a host cell comprising the desired *nonJKL* open reading frames (taken individually or in
20 combination) or *in vitro*, e.g., by providing the polypeptides encoded by the selected *nonJKL* ORFs and the appropriate substrates and/or cofactors. One of skill in the art may also combine the catalytic domains of NonJK, individually or in combination, with other

different catalytic domains so as to form a hybrid enzyme or megasynthetase possessing various catalytic domains useful in synthesizing predicted structures.

In one embodiment, this invention provides for the synthesis of macrotetrolides and/or macrotetrolide analogues or derivatives. In a preferred embodiment, this is accomplished by providing a cell comprising purified *nonJKL* ORFs and culturing the cell under conditions whereby the desired macrotetrolide or macrotetrolide analogue is synthesized. The cell can be a cell that does not normally synthesize a macrotetrolide and *nonJKL* ORFs can be transfected into the cell (i.e., a heterologous host). Alternatively, a cell that typically synthesizes macrotetrolides can be utilized and all or part of the *nonJKL* ORFs can be introduced into the cell. Macrotetrolide derivatives/analogues can be produced by changing the host cell (e.g. to a eukaryotic cell that glycosylates the biosynthetic product), and/or by providing altered metabolites.

The use of standard techniques of molecular biology (gene disruption, gene replacement, gene supplement) can be used to modulate and/or otherwise alter macrotetrolide and/or other metabolite production in an organism that naturally synthesizes a macrotetrolide or an organism that is modified to synthesize a macrotetrolide.

Alternatively, control sequences that alter the expression of various open reading frames can be introduced that alter the amount and/or timing of macrotetrolide production. Thus, for example, by placing particular *nonJKL* open reading frames under control of a constitutive promoters known in the field, macrotetrolide production may be increased.

In addition to the directed modification and/or biosynthesis of various metabolites as described above, the *non/JKL* open reading frames can be utilized, individually or as a module, in combination with other biosynthetic modules (*e.g.* nonribosomal polypeptide synthetases (NRPS) and/or PKS modules and/or enzymatic domains of other PKS and/or NRPS systems) to produce a wide variety of compounds including, but not limited to various macrotetrolides, macrotetrolide derivatives, polyketides, polypeptide derivatives, or polyketide/polypeptide hybrids. General information on modular constructs may be found in Cane et al., Chemistry & Biology, 1999, Vol. 8, No. 12, pps 319-325 and Cane et al., Science, Vol. 282, pps 63-68. The preceding two cited articles are incorporated by reference in their entirety for all purposes. In the present context, the term “module” or “modularity,” refers to coordinated groups of active sites in large, multifunctional proteins in which each coordinated group, or module, is responsible for a particular catalysis step as exemplified by a polypeptide chain elongation as illustrated in Figs. 4-5. Well known modular genetic methodology may be used to alter the number, content, and order of such modules and, in doing so, alter rationally the structure of the resultant products. Entirely unnatural products may thusly be generated.

Target products can be produced, *in vivo* or *in vitro*, by catalytic biosynthesis using combinations of modules from PKSs, NRPSs, and hybrid PKS/NRPS systems, commonly termed megasynthetases. In a preferred embodiment large combinatorial libraries of cells harboring various megasynthetases can be produced by random or directed modification with subsequent selection for the production of a target molecule or molecules of interest. It will be appreciated that, in certain embodiments, such libraries

of megasynthetases, can be used to generate large, complex combinatorial libraries of compounds which themselves can be screened for a desired activity.

In numerous embodiments of this invention, novel molecules, may be created by modifying the *nonJKL* ORFs so as to introduce variations into metabolites synthesized by the enzymatically-active polypeptides. Such variations may be introduced by design, for example to modify a known molecule in a specific way, e.g. by replacing a single monomeric unit within a polymer with another, thereby creating a derivative molecule of predicted structure. Such variations can also be made by adding one or more modules or enzymatic domains to a known PKS, or by removing one or more module from a known PKS.

Furthermore, it is possible to introduce variant nonJKL domains or modules into a megasynthetase. Mutations can be made to the native nonJKL subunit sequences and such mutants used in place of the native sequence, so long as the mutants are able to function with other subunits (domains or modules) in the synthetic pathway. Such mutations can be made to the native sequences using conventional techniques such as by preparing synthetic oligonucleotides including the mutations and inserting the mutated sequence into the gene encoding a subunit using restriction endonuclease digestion. (*see, e.g., Kunkel, (1985) Proc. Natl. Acad. Sci. USA 82: 448; Geisselsoder et al. (1987) BioTechniques 5: 786*). Alternatively, the mutations can be effected using a mismatched primer (generally 10-20 nucleotides in length) which hybridizes to the native nucleotide sequence (generally cDNA corresponding to the RNA sequence), at a temperature below the melting temperature of the mismatched duplex. The primer can be made specific by keeping primer length and base composition within relatively narrow limits and by

keeping the mutant base centrally located (Zoller and Smith (1983) *Meth. Enzymol.* 100: 468). Primer extension is effected using DNA polymerase, the product cloned and clones containing the mutated DNA, derived by segregation of the primer extended strand, selected. Selection can be accomplished using the mutant primer as a hybridization
5 probe. The technique is also applicable for generating multiple point mutations (*see, e.g.,* Dalbie-McFarland *et al.* (1982) *Proc. Natl. Acad. Sci USA* 79:6409). PCR mutagenesis will also find use for effecting the desired mutations.

In another embodiment, variations can be made randomly, for example by making a library of molecular variants by randomly mutating one or more elements of the
10 *nonJKL* genes or by randomly replacing one or more open reading frames of *nonJKL* genes with one or more of alternative open reading frames.

As can be realized from the foregoing, the various open reading frames can be combined into a single multi-modular megasynthetase thereby dramatically increasing the number of possible combinations obtained using these methods. These combinations
15 can be made using standard recombinant or nucleic acid amplification methods, for example by shuffling nucleic acid sequences encoding various modules or enzymatic domains to create novel arrangements of the sequences, analogous to DNA shuffling methods described in Crameri *et al.* (1998) *Nature* 391: 288-291, and in U.S. Patents 5,605,793 and in 5,837,458. In addition, novel combinations can be made *in vitro*, for
20 example by combinatorial synthetic methods. Novel molecules or molecule libraries, can be screened for any specific activity using standard methods.

Random mutagenesis of the nucleotide sequences obtained as described above can be accomplished by several different techniques known in the art, such as by altering

sequences within restriction endonuclease sites, inserting an oligonucleotide linker randomly into a plasmid, by irradiation with X-rays or ultraviolet light, by incorporating incorrect nucleotides during *in vitro* DNA synthesis, by error-prone PCR mutagenesis, by preparing synthetic mutants or by damaging plasmid DNA *in vitro* with chemicals.

- 5 Chemical mutagens include, for example, sodium bisulfite, nitrous acid, hydroxylamine, agents which damage or remove bases thereby preventing normal base-pairing such as hydrazine or formic acid, analogues of nucleotide precursors such as nitrosoguanidine, 5-bromouracil, 2-aminopurine, or acridine intercalating agents such as proflavine, acriflavine, quinacrine, and the like. Generally, plasmid DNA or DNA fragments are
- 10 treated with chemicals, transformed into *E. coli* and propagated as a pool or library of mutant plasmids.

- Large populations of random enzyme variants can be constructed *in vivo* using "recombination-enhanced mutagenesis." This method employs two or more pools of, for example, 10^6 mutants each of the wild-type encoding nucleotide sequence that are
- 15 generated using any convenient mutagenesis technique, described more fully above, and then inserted into cloning vectors.

- In either the directed or random approaches, nucleic acids encoding novel combinations of genes are introduced into a cell. In one embodiment of the present invention, nucleic acids encoding the *nonJ*, *nonK*, *nonJK* or *nonJKL* open reading frames
- 20 are introduced into a cell so as to replace one or more open reading frames of an endogenous gene cluster within a cell. Endogenous gene replacement can be accomplished using standard methods, such as homologous recombination. In a preferred embodiment, such nucleic acids are introduced into the cell optionally along

with a number of additional genes, together called a 'modified gene cluster,' that influence the expression of the genes, survival of the expressing cells, etc. In a particularly preferred embodiment, such cells do not have any other macrotetrolide and/or PKS encoding genes or gene clusters, thereby allowing the straightforward isolation of the molecule(s) synthesized by the genes introduced into the cell.

Furthermore, the recombinant vector(s) can include genes from a single nonactin gene cluster, or may comprise hybrid replacement PKS gene clusters with, *e.g.*, a gene for one cluster replaced by the corresponding gene from another polyketide gene cluster. For example, it has been found that ACPs are readily interchangeable among different synthases without an effect on product structure. Furthermore, a given KR can recognize and reduce polyketide chains of different chain lengths. Accordingly, these genes are freely interchangeable in the constructs described herein. Thus, the replacement clusters of the present invention can be derived from any combination of PKS and gene sets that ultimately function to produce an identifiable polyketide.

Examples of hybrid replacement clusters include, but are not limited to, clusters with genes derived from two or more of the act gene cluster, the *whiE* gene cluster, frenolicin (*fren*), granaticin (*gra*), tetracenomycin (*tcm*), 6-methylsalicylic acid (6-msas), oxytetracycline (*otc*), tetracycline (*tet*), erythromycin (*ery*), griseusin (*gris*), nanaomycin, medermycin, daunorubicin, tylosin, carbomycin, spiramycin, avermectin, monensin, nonactin, curamycin, rifamycin and candidin synthase gene clusters, among others. (For a discussion of various PKSs, *see, e.g.*, Hopwood and Sherman (1990) *Ann. Rev. Genet.* 24: 37-66; O'Hagan (1991) *The Polyketide Metabolites*, Ellis Horwood Limited.

A number of hybrid gene clusters have been constructed, having components derived from the *act*, *fren*, *tcm*, *gris* and *gra* gene clusters (see, e.g., U.S. Patent 5,712,146). Other hybrid gene clusters, as described above, can easily be produced and screened using the disclosure herein, for the production of identifiable polyketides, polypeptides.

Host cells (e.g. *Streptomyces*) can be transformed with one or more vectors, collectively encoding a functional PKS system, or a cocktail comprising a random assortment of polyketide-related ORFs genes, modules, active sites, or portions thereof. The vector(s) can include native or hybrid combinations of such molecules, or mutants thereof. As explained above, the gene cluster need not correspond to a complete native gene cluster but need only encode the necessary PKS system components to catalyze the production of the desired product(s).

In addition to varying the nucleic acids comprising the subject gene cluster, variations in the products produced by the gene cluster(s) can be obtained by varying the host cell, the starter units and/or the extender units. Thus, for example different starter units can be utilized in the synthetic pathway resulting in different polyketide or target molecule variants. Similarly different intermediate metabolites can be provided (e.g. endogenously produced by the host cell, or produced by an introduced heterologous construct, and/or supplied from an exogenous source (e.g. the culture media)). Similarly, varying the host cell can vary the resulting product(s). For example, a gene cassette carrying the *nonJL* biosynthesis genes can be introduced into a deoxysugar-synthesizing host for the production of glycosylated polyketide metabolites.

Before describing an example of the production of unnatural macrotetralides according to the invention, the inventors' foregoing results will be summarized. In brief, the macrotetrolide biosynthetic gene cluster from *Streptomyces griseus* DSM40695 was characterized to encode type II polyketide synthase (PKS) composed of five β -ketoacyl synthases (KSs), four ketoreductase (KRs), two enoyl-coenzyme A (CoA) hydratase (ECH) homologs, and a CoA transferase. The most characteristic feature of the macrotetrolide biosynthetic cluster was the absence of an ACP, which introduced a new type II PKS system working directly on acyl-CoA intermediates. A series of gene inactivation experiments confirmed that all the key components aforementioned play critical roles in the macrotetrolide biosynthesis. The heterologous expression study in *S. lividans* identified a genetic locus responsible for macrotetrolide production, in which two KSs of NonJK catalyze macrotetralization of (\pm)-6-CoA and NonL, a CoA ligase converts (\pm)-6 into (\pm)-6-CoA. A cell-free extract (CFE) study further guided the inventors to define NonJK-catalysis to include NonJ converting 6' into (-)-6-(+)-6-CoA (6'') that is subsequently oligomerized and cyclized into 1 by NonK.

The targeted gene inactivation and *in vivo* (\pm)-6 feeding experiment demonstrated that an ECH, NonS, is specifically involved in the (-)-pathway. The enantiospecific role of NonS was evidenced by accumulation of 2 and 3 in addition to 1 from (\pm)-6-fed SB2003 (*nonS*-inactivation mutant) that no longer produce macrotetrolide, which indicated that the biosynthesis of (+)-7 (the monomeric component of 2 and 3) does not require NonS and whereby, NonS participates specifically in biosynthesis of (-)-6 and (-)-7.

Efficient *in vivo* incorporation of exogenous (\pm)-6 into 1 and the established synthetic method for 6-derivative had presented an opportunity to generate unnatural macrotetrolides though precursor-directed biosynthesis employing a synthetic 6-analogue. To practice this approach, the inventors prepared (\pm)-8-*n*-propyl nonactate (8) that was fed to SB2003 to seek new macrotetrolides derived from this synthetic building block. The feeding, extraction, and chemical procedure was described in Walczak, R. J.; Woo, A. J.; Strohl, W. R.; Priestley, N. D. *FEMS Microbiol. Lett.* 2000, 183, 171. (d) Smith, W.; Xiang, L.; Shen, B. *Antimicrob. Agents Chemother.* 2000, 44, 1809. When 8 was fed to *S. griseus* wild-type, production of new macrotetrolides was evident in HPLC analysis but their identifications and separations were severely compromised by the high abundance of natural macrotetrolides (Figure 16AB). On the contrary, SB2003 provided a clean background due to natural macrotetrolide, by which accumulations of three 8-incorporated macrotetrolides were readily recognizable in HPLC spectrum (Figure 16CD). The yield of new compounds from the 8-fed SB2003 was measured roughly by using authentic 1 (Sigma, St. Louis, MO) as the external standard and it was found that 1 mg of new macrotetrolides was generated from supplementation of 10 mg of 8 into a 50-ml-fermentation. The extract was further analyzed by thin layer chromatography (TLC), in which the compounds were readily separated. Each compound was then isolated and analyzed by HPLC (to correlate TLC to the cognate *peak* in HPLC) and mass spectroscopy (MS) analysis. As annotated in Figure 15, the three compounds showed molecular masses of 875.63 (*m/e* for $C_{47}H_{80}O_{13}Na$, calcd 875.55; 9), 815.54 (*m/e* for $C_{44}H_{72}O_{12}Na$, calcd 815.49; 10), and 843.57 (*m/e* for $C_{46}H_{76}O_{12}Na$, calcd 843.52; 11), respectively, by which 10 and 11 were determined as macrotetrolides composed of two

each of 6 and 8 and one 6 and three of 8, respectively (Figure 11). The monomer compositions of 10 and 11 were confirmed by MS/MS and MS^I analysis that generated daughter molecular ions resulted from the loss of monomeric 6 (*m.w.* 184) or monomeric 8 (*m.w.* 212); fragmentation of 10 resulted in *m/e* 631 and 603 as trimeric structures and *m/e* 419 as dimeric structure. Notably, the heterodimer of 6 and 8 (*m/e* 419) was solely detected as dimeric structure, which was consistent with the expected absence of dimeric 6 (the free acid form of 6'') in SB2003 due to lack of (-)-6. Similarly, 11 was fragmented into *m/e* 659, 631, 447, and 419, verifying the inclusion of 6 and 8. The molecular ion mass of 9 corresponds to the linear tetramer composed of one 7 and three molecules of 8.

10 The compound 9 may be a hydrolyzed product but more plausibly is an intermediate derailed from cyclization. The compound 11 was detected as the most slowly eluted compound in both HPLC and TLC, indicating that formation of other larger macrotetrolides such as 12 (virtual cyclized product of 9) and 13 (homogenous macrotetrolide composed of four molecules of 8) was not permitted by the given specificity of NonK (the dimerization of 8 by NonL/NonJ seemed readily occur as seen in 11).

In the *in vivo* experiment with SB2003, the inventors demonstrated that recruiting endogenous enantiomeric 6 with the exogenous synthetic 6-analogue (8) could generate unnatural macrotetrolides. However, the inventors encountered a limitation in macrotetrolide biosynthetic machinery; that is, 12 and 13 could not be generated. As aforementioned, the inventors established a cell-free system by expressing NonLJK in *S. lividans*. Harnessing this *in vitro* system and the acquired information, previously described, on distinctive roles of NonL, NonJ,

and NonK, 8 was further challenged for formation of unnatural macrotetrolides *in vitro*.

In this particular experiment, the inventors expanded the category of synthetic 8 derivatives to determine the degree of flexibilities of each process in macrotetrolide biosynthesis. With minor modifications in the previously described method, (\pm)-8-*n*-propyl-nonactate (8), (\pm)-*epi*-8-*n*-propyl-nonactate (A), (\pm)-2-*n*-propyl-nonactate (B), and (\pm)-*epi*-2-*n*-propyl-nonactate (C) were prepared. These four compounds were tested for the CoA-activated by using purified NonL. Whereas 8 and A were efficiently converted into their CoA derivatives (8' and A', respectively) by NonL, no conversion was evident with B and C, implicating that introduction of bulky alkyl group near carboxylic group hindered its fitting into substrate pocket of NonL. The dimer formation was further tested with 8 and A by using CFE obtained from *S. lividans* harboring *nonL* and *nonJ*-expression construct [*S. lividans* (pBS2045)], in which 8 was converted the dimeric 8-CoA (8'') but no dimeric A-CoA was found with A, indicating that the correct stereochemistry of hydroxyl nucleophilic C-8 is critical for C-O bond formation. The chromosomal DNA fragment encompassing *nonJ* and *nonL* from *S. griseus* DSM40695 was subcloned into pIJ4123 under the control of thiostrepton inducible promoter (*tipAp*) to generate pBS2045 that was introduced into *S. lividans* TK-24. For pIJ4123, see the reference; Takano, E.; White, J.; Thompson, C. J.; Bibb, M. J. *Gene* 1995, 160, 133.

Compared to the natural substrate such as 6', the dimerization process with 8' was significantly slow but a quantitative conversion was achieved by prolonged incubation (12 h). Identities of 8' and 8'' were verified by MS analysis by detecting characteristic molecular ions (m/e for [M-H]⁺) of 978.222 (C₃₃H₅₆N₇O₁₉P₃S, calcd 978.249) and 1190.363 (C₄₅H₇₅N₇O₂₂P₃S, calcd 1190.390). After defining that 8 could be converted into 8'' *in vitro*, generation of 13 was challenged by incubating 8 with NonLJK-containing CFE that was obtained from *S. lividans* (pBS2048). The pBS2048 was prepared by subcloning *nonL*, *nonJ*, and *nonK* into pWHM467 downstream to *actIp* and *actIIIp*. For pWHM467, see the reference; Wohler, S. E.; Lomovskaya, N.; Kulowski, K.; Fonstein, L.; Occi, J. L.; Gewain, K. M.; MacNeil, D. J.; Hutchinson, C. R. *Chem. Biol.* 2001, 108, 1. The incubation failed to generate 13 even in prolonged incubation, however. The 8' and 8'' were purified and incubated with *S. lividans* (pBS2048)-CFE but 13 could not yet be generated. Thereby, the inventors concluded that 13 can not be generated by NonK, proceeded by *in vitro* generation of 10 and 11. Although 13 was detected as the second major macrotetrolide in the 8-fed SB2118 (the *nonO*-mutant) and at the significant level in the SB2119 (the *nonP*-mutant) and SB2120 (the *nonN*-mutant), the negligible occurrence of 13 in SB2003 posed a reasonable doubt with the identity of 13 that has solely supported by the molecular ion (m/e 871.568) from MS analysis. It seemed possible that some alterations were imposed on stereochemical center(s) before 8 squeezed into 13. We already observed inter-conversions between (+)-6 and (-)-6 and even between 6 and 7 in SB2118 supplemented with (+)-6 or (-)-6.

Furthermore, it was also observed that the same profiles of MTs were found from 8-fed and 8-*epi*-8-fed SB2118, which strongly urged us to consider alteration(s) of 8 to fit into 13.

When 6' and 8' were co-incubated with NonJ, a compound distinguishable from 6'' and 8'' was found as the major product with concomitant occurrence of 6'' (Figure 16-IV). This product was identified as the CoA-derivative of the dimer composed of 8 and 6 (14) by detecting the molecular ion (*m/e* for [M-H]⁺) of 1162.315 (C₄₃H₇₁N₇O₂₂P₃S, calcd 1162.358). With enantiomeric 6 in hand, (-)-6' and (+)-6' were prepared separately and incubated with (±)-8', which showed that (+)-6' was the preferred partner 8' for 14-CoA as (-)-8-(+)-6-CoA (Figure 16-V). Even though significantly limited in the given condition, (-)-6 did generate 14'-CoA as evidenced by MS analysis (Figure 16-VI). The 14'-CoA was supposed to be (-)-6-(+)-8-CoA. The preferred generation of 14-CoA to 6'' suggested the co-incubation of 6' and 8' in *S. lividans* (pBS2048)-CFE. However, the co-incubation led exclusively to 1, which indicated that NonK highly preferred 6'' to 14-CoA and it was apparently required to suppress generation of 6'' to direct synthesis of 10 and 11, as in SB2003. Therefore, (+)-6' and (-)-6' were separately prepared and incubated with (±)-8' in *S. lividans* (pBS2048). Incubation of (+)-6' and (±)-8' resulted in generation of 10 and 11 as observed in MS analysis (Table 1A); it became evident that 1 and 15 were due to the incorporation of (-)-6' in the (+)-6'-preparation as an impurity (the chiral purity of enantiomeric 6 is 97%), which reflected that macrotetrolide to 10 and 11 were much slower than to 6''-derived macrotetrolides. Table 1 illustrates a matrix-

assisted laser desorption ionization (MALDI)-mass spectroscopy (MS) analysis of ethyl acetate extract of *S. lividans* (pBS2048)-cell free extract (CFE) incubation with (+)-6 and (±)-8 (A) and (-)-6 and (±)-8 (B).

Table 1.

	<i>m/e</i> , [M+Na] ⁺	<i>m/e</i> , calcd	rel. int.*	structure**
A	759.414	C ₆₀ H ₆₈ O ₁₂ Na, 759.429	20	1, [4×6] ^c
	787.442	C ₆₂ H ₆₈ O ₁₂ Na, 787.464	100	15, [3×6 + 8] ^c
	815.475	C ₆₄ H ₇₂ O ₁₂ Na, 815.492	44	10, [2×6 + 2×8] ^c
	833.489	C ₆₄ H ₇₆ O ₁₂ Na, 833.502	14	[2×6 + 2×8] ^l
	843.508	C ₆₄ H ₇₈ O ₁₂ Na, 843.523	10	11, [6 + 3×8] ^c
	861.524	C ₆₆ H ₇₈ O ₁₂ Na, 861.534	10	[6 + 3×8] ^l
B	759.391	C ₆₀ H ₆₈ O ₁₂ Na, 759.429	<5	1, [4×6] ^c
	787.413	C ₆₂ H ₆₈ O ₁₂ Na, 787.464	18	16, [3×6 + 8] ^c
	815.441	C ₆₄ H ₇₂ O ₁₂ Na, 815.492	100	17, [2×6 + 2×8] ^c
	833.427	C ₆₄ H ₇₆ O ₁₂ Na, 833.502	<5	[2×6 + 2×8] ^l
		C ₆₄ H ₇₈ O ₁₂ Na, 843.523	n.d. [‡]	
		C ₆₆ H ₇₈ O ₁₂ Na, 861.534	n.d. [‡]	

* the relative intensities in the mass spectra

** the deduced monomeric compositions; ^ccyclized form; ^llinear form (see Figure 11 for the structures)

[‡] not detected or detected but negligible

5 The uncyclized (linear) form of 10 and 11 were detected as significant portions, reflecting they are derailed from NonK. Comparably, incubation of (-)-6' and (±)-8' lead to 17 (an enantiomer of 10) as a predominant macrotetrolide with negligible occurrence of the uncyclized form (Table 1B). The yield of 17 appeared as five times, at least, that of 10 as judged by macrotetrolide analysis but the yield itself might be affected and

10 compromised by occurrence of 1 and/or 15; that is, high abundance of 1 and 15 might suppress the generation of 10. In the line of interpretation, it should also be reminded that 6''-formation from (+)-6'- and (-)-6'-preparation might be dramatically different due to

obvious differences of their roles in 6''-formation. It was observed that a trace of 1 (inevitably, with the intermediacy of 6'') occurred in the incubation with (+)-6'-preparation but not with (-)-6'-preparation, at least within the limit of detection method employed and when *S. lividans* (pBS2048)-CFE was used. The comparable level of 10 and 17, even with limitation in the generation of 14''-CoA (Figure 16-V and VI), and the relatively low level of uncyclized form of 17 assured that NonK-catalysis is the rate-limiting step.

Therefore, the inventors have demonstrated that biosynthesis of unnatural macrotetrolides can be achieved *in vivo* and *in vitro* by direct incorporation of a synthetic analogue. Generation of 10, 11, and 17 *in vitro* verified their identities originally characterized from *in vivo* conversion but was proved not the practical approach considering low conversion yield and a lot of labor in preparation of enantiomeric 6. Comparison of *in vivo* results with SB2003 and those *in vitro* showed that production of enantiomeric 6 by SB2003 was exclusive and in excellent yield.

It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

V. REFERENCES

Ashworth, D. M.; Clark, C. A.; Robinson, J. A. *J. Chem. Soc. Perkin Trans. I* 1989, 1461.

- Ashworth, D. M.; Robinson, J. A. *J. Chem. Soc. Chem. Commun.* 1983, 1327.
- Ashworth, D. M.; Robinson, J. A. *J. Chem. Soc. Perkin Trans. 1* 1988, 1719.
- Ashworth, D. M.; Robinson, J. A.; Turner, D. L. *J. Chem. Soc. Chem. Commun.* 1982, 491.
- 5 Bao, W.; Sheldon, P. J.; Hutchinson, C. R. *Biochemistry* 1999, 38, 9752.
- Bao, W.; Wendt-Pienkowski, E.; Hutchinson, C. R. *Biochemistry* 1998, 37, 8132.
- Begley, T.P.; Kinsland, C.; Taylor, S.; Tandon, M.; Nicewonger, R.; Wu, M.; Chiu, H. -J.; Kelleher, N.; Campobasso, N.; Zhang, Y. *Top. Curr. Chem.* 1998, 195, 93.
- Bibb, M. J.; Findlay, P. R.; Johnson, M. W. *Gene* 1984, 30, 157.
- 10 Bibb, M. J.; White, J.; Ward, J. M.; Janssen, G. R. *Mol. Microbiol.* 1994, 14, 533.
- Birch, A.W.; Robinson, J. A. In *Genetics and Biochemistry of Antibiotic Production*; Vining, L. C., Stuttard, C., Eds.; Butterworth-Heinemann: Boston, 1995; pp 443-476.
- Borrel, M. N.; Pereira, E.; Fiallo, M.; Garnier-Suillerot, A. *Eur. J. Biochem.* 1994, 223, 125.
- 15 Callewaert, D. M.; Radcliff, G.; Tanouchi, T.; Shichi, H. *Immunopharmacology* 1988, 16, 25.
- Carreras, C. W.; Kholsa, C. *Biochemistry* 1998, 37, 2084.
- Choi, K. H.; Kremer, L.; Besra, G. S.; Rock, C. O. *J. Biol. Chem.* 2000, 275, 28201.
- 20 Clark, C. A.; Robinson, J. A. *J. Chem. Soc. Chem. Commun.* 1985, 1568.
- Dreier, J.; Shah, A. N.; Khosla, C. *J. Biol. Chem.* 1999, 274, 25108.
- Dutton, C. J.; Banks, B. J.; Cooper, C. B. *Nat. Prod. Rep.* 1995, 12, 165.

- Ferrer, J. C.; Jez, J. M.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Nat. Struct. Biol.* 1999, 6, 775.
- Fleck, W. F.; Ritzau, M.; Heinze, S.; Gräfe, U. *J. Basic Microbiol.* 1996, 36, 235.
- Fleming, I.; Ghosh, S. K. In *Studies in Natural Products Chemistry*; Atta-ur-
5 Rhaman, Ed.; Elsevier Science: New York, 1996; Vol. 18, pp229-268.
- Funa, N.; Ohnishi, Y.; Fujii, I.; Shibuya, M.; Ebizuka, Y.; Horinouchi, S.
*Nature*1999, 400, 897.
- Hopwood, D. A. *Chem Rev.* 1997, 97, 2465.
- Ishikawa, J.; Hotta, K. *FEMS Microbiol. Lett.* 1999, 174, 251.
- 10 Jez, J. M.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Nat. Struct. Biol.* 2000, 7,
786.
- Jez, J. M.; Bowman, M. E.; Noel, J. P. *Proc. Natl. Acad. Sci. USA* 2002, 99, 5319
and references therein.
- Kilbourn, B. T.; Dunitz, J. D.; Pioda, A. R.; Simon, W. *J. Mol. Biol.* 1967, 30,
15 559.
- Kwon, H. J.; Shen, B. unpublished.
- Kwon, H. J.; Smith, W. C.; Scharon, A. J.; Hwang, S. H.; Kurth, M. J.; Shen, B.
Science 2002, 297, 23.
- Kwon, H. J.; Smith, W. C.; Xiang, L.; Shen, B. *J. Am. Chem. Soc.* 2001, 123,
20 3385; .
- Meadows, E. S.; Khosla, C. *Biochemistry* 2001, 40, 14855.
- Meiners, U.; Cramer, E.; Fröhlich, R.; Webbling, B.; Metz, P. *Eur. J. Org. Chem.*
1998, 2073.

- Metz, P.; Meiner, U.; Cramer, E.; Fröhlich, R.; Wibbeling, G. *Chem. Commun.* 1996, 431.
- Mishira, P. K.; Park, P. K.; Drueckhammer, D. G. *J. Bacteriol.* 2001, 183, 2774.
- Mori, A.; Kaminuma, O.; Ogawa, K.; Nakata, A.; Egan, R. W.; Akiyama, K.;
5 Okudaria, H. *J. Allergy Clin. Immunol.* 2000, 106, S58.
- Mori, A.; Okudaria, H.; Kobayashi, N.; Akiyama, K. *Int. Arch. Allergy Immunol.* 2001, 124, 172.
- Nelson, M. E.; Priestley, N. D. *J. Am. Chem. Soc.* 2002, 124, 2894.
- Nelson, M. E.; Priestley, N. D. *J. Am. Chem. Soc.* 2002, 124, 2894.
- 10 Pan, H.; Tsai, S.; Meadows, E. S.; Miercke, L. J.; Keatinge-Clay, A. T.;
O'Connell, J.; Khosla, C.; Stroud, R. M. *Structure* 2002, 10, 1559.
- Plater, R.; Robinson, J. A. *Gene* 1992, 112, 117.
- Rawling, B. J. *Nat. Prod. Rep.* 2001, 18, 231.
- Shen, B. *Top. Curr. Chem.* 2000, 209, 1.
- 15 Shen, B.; Kwon, H. J. *Chem. Rec.* in press.
- Smith, W.; Xiang, L.; Shen, B. *Antimicrob. Agents Chemother.* 2000, 44, 1809.
- Spavold, Z. M.; Robinson, J. A. *J. Chem. Soc. Chem. Commun.* 1988, 4.
- Stahl, P.; Pape, H. *Arch. Mikrobiol.* 1972, 85, 239.
- Staunton, J.; Wilkinson, B. *Top. Curr. Chem.* 1998, 195, 49.
- 20 Summers, R. G.; Ali, A.; Shen, B.; Wessel, W. A.; Hutchinson, C. R.
Biochemistry 1995, 34, 9389.
- Takano, E.; White, J.; Thompson, C. J.; Bibb, M. J. *Gene* 1995, 160, 133.
- Tanouchi, Y.; Shichi, H. *Immunology* 1988, 63, 471.

- Teunissen, M. B. M.; Pistor, F. H. M.; Rongen, H. A. H.; Kapsenberg, M. L.; Bos, J. D. *Transplantation* 1992, 53, 875.
- Tsai, S. C.; Miercke, L. J.; Krucinski, J.; Gokhale, R.; Chen, J. C.; Foster, P. G.; Canc, D. E.; Khosla, C.; Stroud, R. M. *Proc. Natl. Acad. Sci. USA* 2001, 18, 14808.
- 5 Walczak, R. J.; Woo, A. J.; Strohl, W. R.; Priestley, N. D. *FEMS Microbiol. Lett.* 2000, 183, 171.
- Wang, Y.; Metz, P. *Tetrahedron: Asymmetry* 2000, 11, 3995.
- Watanabe, C. M. H.; Townsend, C. A. *Chem. Biol.* 2002, 9, 981.
- Wohlert, S. E.; Lomovskaya, N.; Kulowski, K.; Fonstein, L.; Occi, J. L.; Gewain, K. M.; MacNeil, D. J.; Hutchinson, C. R. *Chem. Biol.* 2001, 108, 1.
- 10 Woo, A. J.; Strohl, W. R.; Priestley, N. D. *Antimicrob. Agents Chemother.* 1999, 43, 1662.
- Zawada, R. J. X.; Khosla, C. *Chem. Biol.* 1999, 6, 607.
- Zizka, Z. *Folia Microbiol. (Prague)* 1998, 43, 7.